Aalto University
School of Business

QUANTIFYING PROMOTION EFFECTIVENESS FOR A RETAILER

Master's Thesis
Aleksi Pesonen
Aalto University School of Business
Information and Service Economy
Spring 2018

Aalto University
School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Aleksi Pesonen

Title of thesis Quantifying promotion effectiveness for a retailer

Degree Master of Science in Economics and Business Administration

Degree programme Information and Service Management

Thesis advisor(s) Pekka Malo

Year of approval 2018 **Number of pages** 73 **Language** English

Abstract

Retailers spend a substantial part of their marketing budgets on promotions. While promotions already account for a large share of retailers' revenues, promotion intensity continues to increase. Despite substantial investments in promotions, companies often have little understanding of the true performance and thus struggle to determine which promotions are working and which are not.

The objective of this thesis is to quantify the impact of sales promotions, which can help retailers to identify the most and the least effective promotions. This thesis addresses the research problem by seeking answers to three key questions. First, it studies whether promotions generate a lift in sales during the promotional period. Second, the extent to which the sales lift may cannibalize the sales of other non-promoted products in that category is determined. Lastly, this thesis investigates whether promotions affect the quantities sold in the weeks after the promotion.

The promotion effects were quantified using a 2-year dataset of weekly point-of-sales (POS) data obtained from a Finnish retailer. During the study period, over 140,000 promotions were implemented in nearly 500 different product categories.

The main findings of this thesis are threefold. First, about three quarters of promotions generate additional sales. The magnitude of this sales lift is about 130% in comparison to normal level of sales. The interquartile range of the sales lift measure falls between 17% and 381% indicating high variation among products. Second, approximately 40% of the sales lift accounts to switching from other non-promoted products in the same category. This indicates that on average only about 60% of the additional sales are truly incremental to retailers. Lastly, this thesis did not find evidence for stockpiling effects which refers to decreased quantities sold in the weeks after the promotion period.

The results imply that not all promotions generate the desired results and that the effectiveness of promotions is significantly diluted due to switching. This means that retailers should focus promotional activities to products and categories that are more effective and generate a positive impact. Systematic analysis based on quantitative data applied in this thesis helps retailers to gain insights from a large set of data and understand the true performance of promotions while guiding effective promotional decisions in the future.

Keywords retailer promotions, promotion effectiveness, retail

Aalto-yliopisto
Kauppakorkeakoulu

Aalto-yliopisto, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Maisterintutkinnon tutkielman tiivistelmä

Tekijä Aleksi Pesonen

Työn nimi Myyntikampanjoiden vaikutuksien kvantifiointi vähittäiskaupan näkökulmasta

Tutkinto Kauppatieteiden maisteri

Koulutusohjelma Information and Service Management

Työn ohjaaja(t) Pekka Malo

Hyväksymisvuosi 2018 **Sivumäärä** 73 **Kieli** Englanti

Tiivistelmä

Merkittävä osa vähittäiskauppojen markkinointibudjeteista käytetään myyntikampanjoihin. Vaikka myyntikampanjat vastaavat jo suurta osaa vähittäiskauppojen kokonaisliikevaihdosta, niitä käytetään yhä tiheämmin. Tuntuvista panostuksista huolimatta kampanjoiden todellisen vaikuttavuuden arviointi on hankalaa, ja yrityksille jää usein epäselväksi, mitkä kampanjat ovat onnistuneet ja mitkä eivät.

Tämän tutkimuksen tavoitteena on kvantifioida myyntikampanjoiden vaikutukset, minkä avulla vähittäiskaupat pystyvät tunnistamaan tehokkaat ja tehottomat myyntikampanjat. Tämä tutkielma tarkastelee aihetta kolme avainkysymyksen kautta, joiden avulla kampanjoiden tehokkuutta voidaan arvioida. Ensiksi tutkielmassa määritellään saavatko myyntikampanjat aikaan lisämyyntiä suhteessa normaalitasoon. Tämän jälkeen tutkitaan missä määrin saavutettu lisämyynti kannibalisoi muiden ei-kampanjatuotteiden myyntiä tuotekategorian sisällä. Lopuksi tutkielmassa selvitetään vaikuttavatko myyntikampanjat tulevien viikkojen myyntimääriin.

Myyntikampanjoiden vaikutuksien kvantifiointiin sovellettiin suomalaisen vähittäiskauppaketjun viikottaisia myyntitietoja kahden vuoden ajalta. Tutkimusjakson aikana toteutettiin yli 140 000 kampanjaa lähes 500 eri tuotekategoriassa.

Tämän tutkielman tärkeimmät löydökset voidaan jakaa kolmeen osaan. Ensinnäkin kolme neljäsosaa myyntikampanjoista saa aikaan lisämyyntiä. Myynnin normaalitasoon verrattaessa lisämyynnin suuruus on noin 130 %. Lisämyynnin kvartiiliväli asettuu 17 %:n ja 381 %:n välille, mikä kuvastaa suurta vaihtelua tuotteiden välillä. Toiseksi, noin 40 % lisämyynnistä johtuu ei-kampanjatuotteiden myynnin siirtymisestä tuotekategorian sisällä. Tämä tarkoittaa käytännössä sitä, että kampanjan aikaansaamasta lisämyynnistä vain 60 % tuo aidosti lisämyyntiä. Lopuksi tutkielmassa selvitettiin vaikuttavatko myyntikampanjat tulevien viikkojen myyntimääriin, mutta tällaista yhteyttä ei löydetty.

Tämän tutkimuksen mukaan kaikki myyntikampanjat eivät saa toivottua vaikutusta aikaan. Lisäksi myynnin siirtymistä tuotekategorian sisällä tapahtuu paljon, mikä heikentää kampanjan tehokkuutta merkittävästi. Tutkimustulokset osoittavat, että vähittäiskauppojen tulisi kohdistaa myyntikampanjoita tuotteisiin ja kategorioihin, jotka ovat tehokkaampia ja saavat aikaan myönteisiä vaikutuksia. Tutkielmassa sovellettu systemaattinen analyysi kvantitatiiviseen tietoon perustuen auttaa vähittäiskauppiaita ymmärtämään kampanjoiden todellisia vaikutuksia ja tukee tehokasta päätöksentekoa jatkossa.

Avainsanat myyntikampanja, myyntikampanjoiden tehokkuus, vähittäiskauppa

Table of Contents

List of Tables

List of Figures

1 Introduction

In retailing, companies spend a substantial part of their marketing budgets on promotions that form a particularly attractive tool for goal-oriented managers seeking fast returns on their marketing investments (Ailawadi, Harlam, Cesar, & Trounce, 2006; Van Heerde & Neslin, 2017). While promotions already account for 10 to 45 percent of brick-and-mortar retailers' total revenues, the promotion intensity continues to increase with promotional spending growing at a higher rate than sales volume (Boston Consulting Group, 2015; Dipl-Kfm, 2015; McKinsey&Company, 2013b).

Despite substantial investments in promotions, it remains unclear for many retailers whether the net effect of promotions is positive, not only on profit, but even on sales (Ailawadi & Gupta, 2014). One of the typical pitfalls that hinder the success of promotions is that retailers are having little understanding of true performance (Boston Consulting Group, 2015). All too often plans rely on what was done previous year rather than systematic analysis of historical performance, which sometimes end up hurting instead of helping sales (Boston Consulting Group, 2014; Boston Consulting Group, 2015; Boston Consulting Group, 2017). To stay competitive, retailers need to ensure effective promotional decisions. Therefore, determining which promotions are working and which are not remains an ongoing concern for many, as they are striving to improve effectiveness and increase accountability in their promotional activities. Retailers often seek answers to questions like: Do promotions benefit the company from a managerial perspective? Which promotions generate incremental sales to the company? Are some categories better fit for promotions than the others are?

Studies suggest that many companies base their promotional strategies on theories rather than data and they struggle taking advantage of all the information around transactions that could make a difference in sales (Boston Consulting Group, 2014; McKinsey&Company, 2013b). In fact, many retailers do have the data, but all too often, they are not able to look back and learn from it (Boston Consulting Group, 2014). It is common for retailers to evaluate only actual sales from a promotion rather than incremental sales, implicitly stating that promotion decisions are based only on partial information (Boston Consulting Group, 2015).

Using data effectively and employing analytics, retailers can significantly improve their promotion performance, especially in promotion driven categories (Boston Consulting Group, 2017; McKinsey&Company, 2013b). Translating systematic analysis into concrete actions may not be easy but surely can help retailers to make practical decisions faster and easier. Thus, an in-depth analysis of promotion effects is a necessary starting point for improving effectiveness of promotions in the future.

Considerable amount of research has been focusing on promotional impact from the manufacturer perspective (Ailawadi et al., 2006). While the dynamics of sales promotions apply equally for both retailers and manufacturers, the impact of promotions have distinct managerial implications for each (Van Heerde & Neslin, 2017). This thesis contributes to existing literature by studying promotion effects from a retailer’s perspective that so far has been researched to a lesser extent.

1.1 Objectives and research questions

The ultimate objective of this thesis is to extract information from a large set of data to improve promotional decision making for a retailer. More specifically, this thesis will focus on the impact of sales promotions on unit sales. Research on this topic helps retailers to identify the most and least effective promotions and plan better their future promotion decisions.

The concrete aims of the theoretical part are two-fold. Firstly, the theoretical part provides a solid foundation for understanding the key issues in sales promotion literature that are relevant for this thesis. Secondly, it will introduce common approaches derived from the literature that apply to the research problem at hand. The empirical part of the thesis focuses on the quantitative analysis of the historical point-of-sales (POS) data obtained from a retailing company.

Despite the prevalence of promotions in retailing, implementing effective promotion management remains a challenge (Dipl-Kfm, 2015). According to The Boston Consulting Group (2015), 20 to 50 percent of promotions do not generate noticeable lift in sales or, worse, have a negative impact. Moreover, an academic study on promotion effectiveness in retailing showed that more than 50% of all promotions were not profitable for the retailer (Ailawadi, Harlam, César, & Trounce, 2007). To manage promotions effectively, one needs to evaluate and quantify the effects that promotions have on sales.

Based on the objectives and research problem, the specific research questions are as follows:

1. **Do promotions generate a lift in sales during the promotional period?**

2. **To what extent is the sales lift switched from other non-promoted products within a category?**

3. **Do promotions affect future sales via stockpiling?**

The first two research questions focus on the impact of promotion during the promotional period, which is the main focus of this thesis. The third research question extends the first two and relates to the impact of a promotion beyond the promotional period by raising the question whether promotions affect future sales.

This thesis contributes to both academics and practitioners by creating a better understanding of promotion effects and developing a systematic approach for analysing promotions using quantitative data. Quantifying promotion effects can help retailers to distinguish the most and the least effective promotions and to act accordingly.

1.2 Main findings

The main findings of this thesis are threefold. First, approximately three quarters of promotions generate a lift in sales. The median sales lift found in this study suggests approximately 130% increase in unit sales in comparison to baseline sales which describe the “normal” level of sales in that week if the promotion had not been run. Second, approximately 40% of this sales lift happens due to switching from other non-promoted products in a category. This means that only a portion of the sales lift generated by a promotion is incremental for the retailer. Lastly, stockpiling effects were studied and no significant relationship was found between the sales lift and the category sales in the weeks after the promotion.

1.3 Structure

The rest of this thesis is structured as follows. Chapter 2 Literature review intends to familiarize the reader with how the sales promotions have been studied previously and how this study positions itself within the discipline. Chapter 3 on methods and data explains carefully how the analysis was conducted, characterizes the data and presents the methods used for modelling the promotion effects. Chapter 4 Empirical analysis and discussion describes the quantitative analysis in detail. It starts with describing the data preparation steps needed prior to fitting the statistical models. Then it systematically answers each research questions laid out in this chapter while discussing the main findings. In Chapter 5 the conclusions will be presented. The last chapter also discusses the limitations of the study and gives direction for future research.

2 Literature review

Both academics and practitioners have been intrigued by sales promotions over the past decades and plenty of research have been conducted regarding sales promotions. However, there are still several drivers fostering research in the field of sales promotions. As discussed, substantial investments are allocated to promotions each year with no expected changes in the increasing promotional spending. Secondly, a part of this trend can be attributed to prevalence of POS data that allows modelling and optimizing promotions better and more accurately than previously (Blattberg & Briesch, 2012). Furthermore, recent trends towards big data and analytics are likely to accelerate interest in promotions even further.

Literature on sales promotions is vast and scattered, as many researchers have worked in this field across decades using wide-ranging terminologies, methodologies, and data. Therefore, the purpose of this chapter is to acquaint the reader with how sales promotions have been studied and how this study positions itself within the discipline.

2.1 Introduction to sales promotions

Blattberg & Neslin (1990) define sales promotion as an action-focused event with an objective to affect directly the behaviour of customer. Typically, promotions are temporary and call-to-action (Blattberg & Briesch, 2012). Simply put, sales promotion refers to short-term incentive to encourage purchase or sales of a product or service (Kotler & Armstrong, 2014). What sets promotions apart from advertising is that sales promotions give a direct incentive to buy whereas advertising presents with a reason to buy.

2.1.1 Types of promotions

There are three main types of promotions that are designed for different purposes and different target audiences (Blattberg & Briesch, 2012). Manufacturers may offer *trade promotions* to members of their distribution channel with the objective to stimulate retailers passing through promotions to consumers. When a retailer offers a promotion to consumers, it refers to a *retailer promotion*. The third type of promotion is the one where manufacturers target consumers directly with *consumer promotions*. Figure 1 illustrates the interrelationship between different types of promotions and stakeholders. This thesis concentrates solely on retailer promotions. In the remainder of this paper, promotion will refer particularly to retail promotions. Hence, trade promotions and consumer promotions are left beyond the scope of this study.

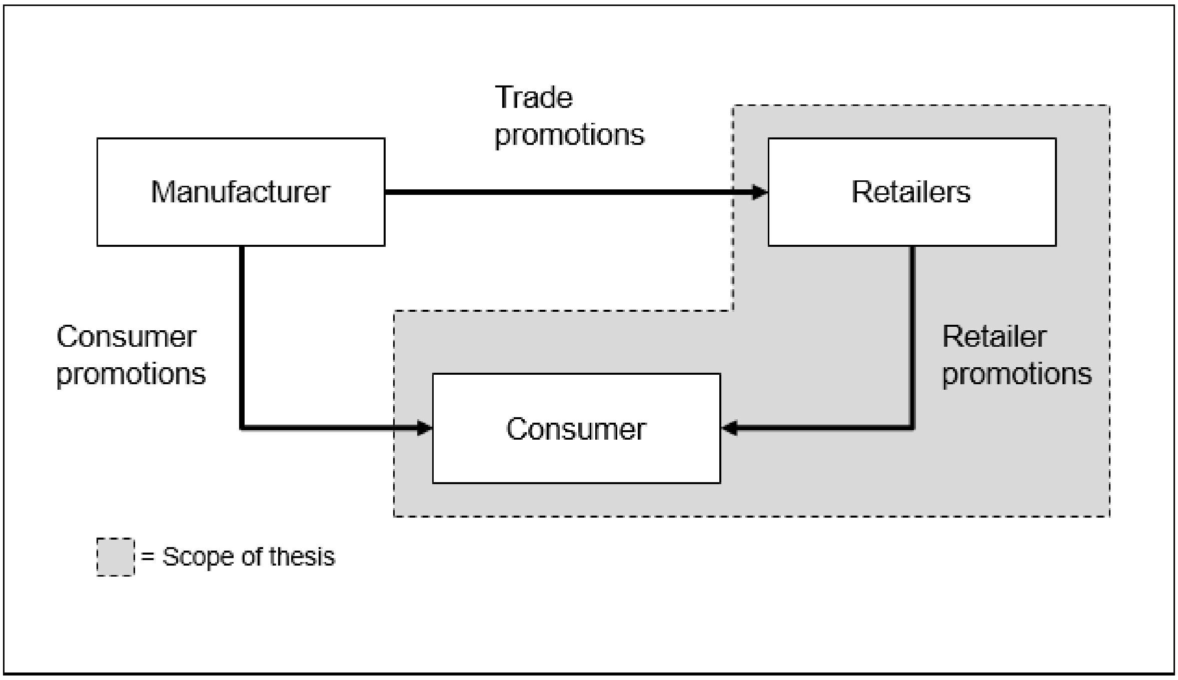

Figure 1. Interrelationship between different types of promotions. Adapted from Blattberg & Briesch (2012)

One can further categorize retailer promotions into various forms. The most common type of retailer promotion is a price discount in which retailers temporarily decrease price of a product. Other forms of retailer promotions include, among others, buy-one-get-one free

(BOGO) promotions and bundle promotions. Table 1 summarizes the common types or forms of retailer promotions and their descriptions.

Table 1: Common retailer promotions. Adapted from Blattberg & Briesch (2012)

Retail promotion	Description
Bundle	Discount given for purchasing products from complementary categories
Coupon	Issued coupons for a product in an advertisement or on the shelf
Free goods	Goods themselves serve as the discount. Includes BOGOs (or buy X, get Y free) and promotions where products in complementary categories are given away
Free trial	Free samples given away to encourage purchase of a new product
N-for	Discounted price for the purchase of a set number of items (e.g. 3 for 5 €)
Price reduction	Temporary price reduction on a product

Various types of promotions may have different levels of effectiveness and induce different types of behaviour. However, this thesis does not try to distinguish between different types of promotions. Instead, the focal point is how promotion, regardless of type, affects sales. Therefore, investigation of the most effective promotions among the various types is deliberately left out of the scope and for other researchers to explore.

2.1.2 Taxonomy of promotion effects on sales

Promotions can induce many different effects on sales of which evaluation remains a continuous challenge for retailers (Ailawadi & Gupta, 2014; Gedenk, Neslin, & Ailawadi, 2010). The immediate sales lift i.e. gross lift of a promotion (or more commonly known as the “sales promotion bump”) refers to the portion of sales of that is incremental compared to the sales if no promotion was in place. The estimated sales that would have occurred without running the promotion is referred to baseline sales. Consequently, sales promotion bump is the difference between actual sales during the promotion and the baseline sales.

Increases in sales can result in many different effects during and after the promotional period. In order to assess the effectiveness of a promotion, one needs to untangle these effects, since the effectiveness depends not only on the immediate sales lift but also on the sources of this lift (Ailawadi et al., 2006; Gedenk et al., 2010; Van Heerde & Neslin, 2017). Consistent with Gedenk et al. (2010) short-term effects is defined in this thesis as effects that occur during promotional period whereas long-term impact refers to effects that occur after the promotional period.

During the promotional period, promotions can for example induce new users to buy the promoted item or existing users to consume more. On the other hand, the promotions may result in consumers switching between brands, categories or stores. In addition to short-term effects, promotion may have long-term effects on sales after the promotional period. For instance, when consumer stockpile extra quantity for future use, this increases sales during the promotion but decreases them afterwards (Gedenk et al., 2010). Another long-term effect is customer loyalty which may change due to promotion. Manufacturers naturally wish for high brand loyalty whereas retailers may be more interested in store loyalty. Taxonomy of different effects on sales of a promoted product discussed above is illustrated in Figure 2.

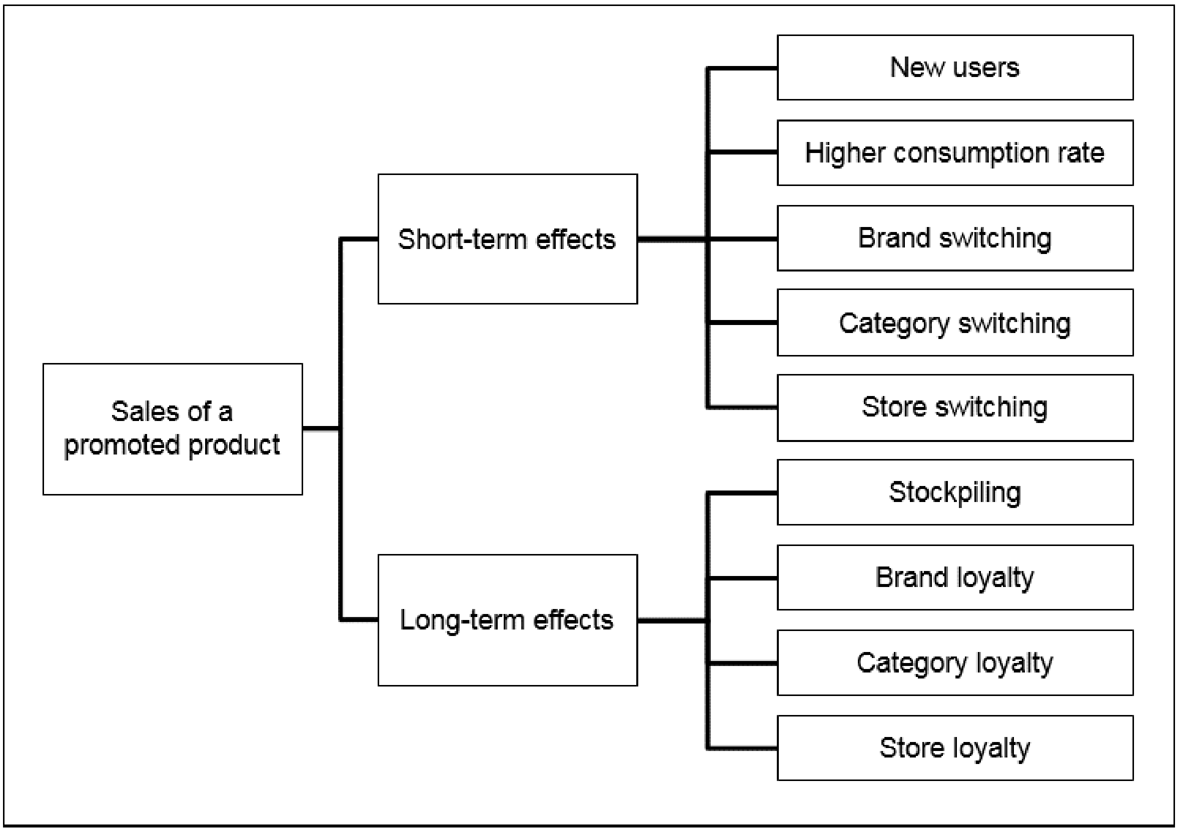

Figure 2. Taxonomy of sales promotion effects. Adapted from Gedenk et al. (2010). For a more detailed list of effects with a slightly different categorization, see Van Heerde & Neslin (2017), p. 16-17

A related term that is not included in the Figure 2 but often found in the literature when talking about promotion effects is purchase acceleration. Purchase acceleration comprises of two parts: timing acceleration and quantity acceleration. Timing acceleration means that consumers buy earlier due to promotion whereas quantity acceleration refers to consumers buying more. Hence, purchase acceleration can result in either increased consumption or stockpiling extra quantity for future use (Gedenk et al. 2010). Another related concept that is not included in Figure 2 is the distinction between primary and secondary demand effects of promotion. Primary demand refers to cross-period effects (e.g. stockpiling) and category expansion (e.g. increased consumption) whereas secondary demand effects refer to brand switching.

2.1.3 Distinct perspectives by retailers and manufacturers

It is important to understand that manufacturers and retailers often pursue quite different objectives with promotions. Commonly, retailers are interested in increasing the sales in their stores of the product category to which a product belongs, not just the product alone. This means that retailers need to comprehend category dynamics of promotions to design effectively their promotional strategies (Abraham & Lodish, 1993; Blattberg & Briesch, 2012).

In contrast to the retailer, however, the main objective of manufacturer is to increase the sales of a promoted product. For example, switching from competitors' products to the promoted item is naturally beneficial for the manufacturer. However, this is not necessarily true for a retailer if one switches consumption merely within the product category and category sales do not increase altogether. If this is the case, then it is likely to cause a negative impact for a retailer because switching happens from non-promotional items to a promoted item that typically has lower margin. Therefore, the total category sales are oftentimes more interesting for retailers rather than individual product sales. Table 2 puts together selected unit sales effects from the vantage points of retailers and manufacturers.

Table 2: Unit sales effects of promotions for retailers and manufacturers. Adapted from Van Heerde & Neslin (2017)

Source	Retailer	Manufacturer
New users	positive	positive
Consumption rate	positive	positive
Store switching	positive	neutral
Brand switching	neutral	positive
Category switching	neutral	positive or neutral(1)
Stockpiling	neutral	neutral

(1) If manufacturer produces product in the other category that is substituted, it does not benefit from the category switch (neutral effect), otherwise it has a positive effect on unit sales

Note: Positive effect denotes that a party benefits from the source of deal volume.

Neutral effect denotes that the party does not benefit from the source of deal volume

2.2 Decomposition of promotional effects

Many researchers have been interested in the sources of sales lift in studying the effects of promotions. There are two fundamental approaches for decomposition of sales promotions effects: the "elasticity approach" and "unit sales approach" (Van Heerde & Neslin, 2017). Some papers may also refer these to "gross" and "net" approaches, respectively (Van Heerde, Leeflang, & Wittink, 2004). The former definition is more commonly used and therefore employed throughout this thesis.

2.2.1 Elasticity approach

Elasticity approach contains customer-level decision models based on household data. It is grounded on the perspective that promotions can affect purchase incidence, brand choice and purchase quantity decisions. These consumer decisions have received the most attention in earlier literature (Van Heerde & Neslin, 2017).

Elasticity approach was originated by Gupta in 1988. His seminal paper distinguishes three components of household response to a promotion: category purchase timing, brand choice and purchase quantity (Van Heerde, Gupta, & Wittink, 2003). Gupta decomposes the sales promotion bump into sales lift due to brand switching, purchase time acceleration and stockpiling that aid in understanding effectiveness of promotions. Elasticity model assumes that brand choice and purchase quantity are conditional on the purchase time decision (Gupta, 1988). More specifically, probability that a household *h* buys *q* units of brand *b* during week *t* is the product of three probabilities (Van Heerde & Neslin, 2017):

1) Probability that household *h* buys the category in week t (incidence)
2) Probability that, conditional on incidence at time t, household *h* buys brand *b* (choice)
3) Probability that, conditional on incidence and a choice to buy brand *b* during week *t*, the household *h* buys *q* units (quantity)

Consumer decisions regarding purchase incidence, choice and quantity are often modelled separately using different methods (Van Heerde & Neslin, 2017). For example, binary logit models are commonly used to model category incidence (Bucklin, Gupta, & Siddarth, 1998), multinomial logit models are applied to brand choice (Guadagni & Little, 1983) whereas Poisson models are used to model purchase quantity (Bucklin et al., 1998). Once each decision model has been estimated, elasticities can be calculated for each.

Van Heerde & Neslin (2017) illustrate the decomposition well with an example: assume a total sales elasticity of -3 with respect to promotion is calculated and could be decomposed as -3 (sales elasticity) = -0.45 (incidence) -2.25 (choice) -0.3 (quantity). That is, incidence elasticity would comprise 15%, brand switching 75% and quantity elasticity 10% of the total sales elasticity. Hence, the foundation of the elasticity approach relies on the mathematical relationship that the elasticity of the probability of buying brand b at shopping trip t with respect to promotion is equal to the sum of elasticities of purchase incidence, brand choice and purchase quantity with respect to promotion (Van Heerde & Neslin, 2017). In simple terms, elasticity approach tries to explain how promotions can influence whether or when consumers buy (purchase incidence), what they buy (brand choice), and how much (purchase quantity) by decomposing an item's total price elasticity into these components.

Gupta (1988) was first one to study all three consumer decisions in a single framework albeit each were modelled separately (Chintagunta, 1993). Elasticity approach has been researched widely since and extended by many academics (Bucklin et al., 1998; Chiang, 1991; Chintagunta, 1993). Some models extend Gupta's model by integrating two or more consumer decisions (Chiang, 1991; Chintagunta, 1993). For example, model proposed by Chiang (1991) pose the key questions together as an integrated model (whether, what and how much to buy) within the context of a single utility maximization problem. Interestingly, this integrated model produced practically identical results (see Table 3 in the end of this section) to Gupta's non-integrated model.

Other models have augmented elasticity approach to account for consumer heterogeneity. This type of extension stems from the assumption that consumers are naturally heterogeneous in their brand preferences, responsiveness to promotions and other marketing activities as well as how much they learn from product usage experience (Van

Heerde & Neslin, 2017). Consequently, one should allow parameters in the incidence, choice and quantity models to vary across customers.

Furthermore, there are vast amount of other extensions to the basic elasticity approach that focus on different aspects of sales promotion impact. These include, among others, store switching (Bucklin & Lattin, 1992), category switching (Manchanda, Ansari, & Gupta, 1999), cannibalization (Fader & Hardie, 1996) and deceleration (Macé & Neslin, 2004; Van Heerde, Leeflang, & Wittink, 2000). Notably, while these customer or household level models are largely used in many decomposition studies, one could also use similar models to for studying other issues as well, such as brand loyalty (Van Heerde & Neslin, 2017).

Overview of empirical results based on the elasticity approach is presented in Table 3. Empirical generalizations based on elasticity decomposition suggest that the majority of promotion effects, approximately 70% on average, stem from consumers switching from competing products.

Table 3: Decomposition results from elasticity studies. Adapted from Van Heerde & Neslin (2017)

Study	**Category**	**% Secondary demand effects (switching)**	**% Primary demand effects**
Gupta (1988)	Coffee	84	16
Chiang (1991)(1)	Coffee	83	17
Chintagunta (1993)	Yogurt	40	60
Bucklin et al. (1998)	Yogurt	58	42
Bell et al. (1999)(2)	Multiple categories	75	27
Van Heerde et al. (2003)(3)	Multiple categories	57	42
Chib et al. (2004)(4)	Cola	70	30
Nair et al. (2005)	Orange juice	65	35
Average elasticity decomposition		***67***	***34***

(1) Average results from two types of promotions in coffee category
(2) Average results across 13 different categories
(3) Average results across three categories
(4) Average results across three types of promotions
Note: effects may not sum up to 100% due to rounding

2.2.2 Unit sales approach

Absolute sales effects refer to changes in unit sales as a result of a sales promotion. Since promotions are, by definition, intended to promote sales, absolute sales effects in comparison to elasticities are preferable in measuring the effectiveness of sales promotions (Van Heerde, 2005). The origins of unit sales approach stem from early 2000s when a complementary decomposition measure for promotion impact was develop based on unit sales (Van Heerde et al., 2003).

There are different ways to go about modelling unit sales effects, but the central idea is to focus on changes in actual sales of promoted products as well as other products within a category, instead of estimating probabilities and their elasticities for consumer decisions. In comparison to elasticity studies, unit sales approach uses store-level data and accounts for one of the key managerial questions regarding promotions: "If promoted brand gains 100 units, how many units other brands lose, how many units come from other time periods and how many units represent category expansion" (Van Heerde et al., 2004). In their paper, Van Heerde et al. (2003) derive analytical expressions that transform elasticity decomposition into unit sales effect to demonstrate the differences between the approaches. Their key finding was that studies based on elasticity decompositions have been interpreted in an inadequate way. Justification for this will be elaborated in section 2.2.3 where empirical generalizations from elasticity and unit sales studies are discussed in detail.

Sales promotion literature that is more recent has relied on the unit sales approach. For example, to determine whether sales lift is truly beneficial from a managerial perspective, Van Heerde et al. (2004) developed a regression-based method for decomposing sales effects into sales taken from competing products in the promotion week (switching), sales taken from other time periods and total increases in category sales. In addition, the authors provide two extensions to the basic model that separate category expansion effects into cross-store and market-expansion effect and cross-item effects into cannibalization and between-brand effects both of which are often in the interest of manufacturers.

Studying promotion effects in terms of unit sales allows measuring profit impact of promotions. Ailawadi et al. (2006, 2007) are the only ones that have quantified profit impact of individual promotions. Commonly, measuring profit impact is often out of the question due to lack of publicly available cost data, which may explain the lack of published papers. Part of this challenge can also be attributed to difficulties in allocating costs and trade allowances for individual promotions that is a prerequisite for calculating promotional margins accurately. In their extensive and recognized paper, Ailawadi et al. (2007) quantify net unit and net profit impact of individual promotions for a drug store retail chain applying unit sales approach in estimating the promotion effects. They compiled data from all promotions offered by the retail chain over a one-year period and estimated the gross lift of the promotion, decomposed the gross lift into three components (switching, stockpiling, category expansion), and estimated the extent to which promotion increases sales of other product categories in the store. Furthermore, they conducted a meta-analysis on the correlates of the impact and interestingly found that many promotion and brand characteristics have opposing associations with the net unit and profit impact.

Other studies have applied unit sales approach as well. For example, in a study about primary and secondary demands effects with aggregate data (Nair, Dubé, & Chintagunta, 2005), the authors use both elasticity-based (Bell, Chiang, & Padmanabhan, 1999) and unit sales approach (Van Heerde et al., 2003) to decompose price elasticity of demand and predicted sales into switching and purchase incidence / quantity effects. They contribute to existing literature by developing a method for estimating the aggregate demand system corresponding to the models of consumer choice of Chiang (1991) and Chintagunta's (1993). Unit sales approach has also been used in studying how promotions affect endogenous consumption (Sun, 2005). Sun uses unit-sales based approach to break down promotional sales increase into switching, stockpiling and consumption increase in two product categories.

Table 4 presents an overview of research results based on unit sales decomposition. Studies based unit sales approach suggest that switching from competing products (secondary demand effect) accounts for much smaller portion of the sales lift than was previously thought based on elasticity studies.

Table 4: Decomposition results from unit sales studies. Adapted from Van Heerde & Neslin (2017)

Study	Category	% Secondary demand effects (switching)	% Primary demand effects
Pauwels et al. (2002)(1)	Multiple categories	25	75
Van Heerde et al. (2003)(2)	Multiple categories	33	67
Sun et al. (2003)	Ketchup	56	44
Van Heerde et al. (2004)(3)	Multiple categories	32	69
Nair et al. (2005)	Orange juice	8	92
Sun (2005)(4)	Multiple categories	41	60
Chan et al. (2008)(5)	Multiple categories	21	79
Leeflang et al. (2008)(6)	Multiple categories	20	80
Average unit sales decomposition		***29***	***71***

(1) Average results across two categories
(2) Average results across three categories
(3) Average results across four categories
(4) Average results across two categories
(5) Average results across 2 categories
(6) Average results across 7 categories
Note: effects may not sum up to 100% due to rounding

2.2.3 Empirical generalizations based on elasticity versus unit sales approach

Research based on elasticity approach relying on household panel data suggest that majority of the incremental promotional sales follow from switching from competing products. Van Heerde & Neslin (2017) reviewed that on average about 70% of sales promotion elasticity is attributed to switching and the remainder to timing acceleration and quantity effects (i.e. primary demand effects). From Table 3 we may observe that there is a downward trend in switching percentages based on elasticity studies. For example, Gupta (1988) and Chiang (1991) report that over 80% of the incremental sales comes from switching. Other studies suggest slightly lower percentages in the range of 40-75%. In comparison, from Table 4 we see that corresponding switching percentages based on unit sales approach yields much lower switching effects, in most cases between 20-40%.

Several academics have interpreted based on elasticity studies that if product gains 100 units in sales during a promotion the other products in the category lose, on average, about

75 units (Van Heerde et al., 2003). In their study, Van Heerde et al. (2003) demonstrate that this interpretation is incorrect due to way the two approaches treat category expansion that occurs during a promotion. The authors show that differences in the effects arise from the fact that elasticity decomposition studies do not account for category growth induced by the increase in purchase incidence probability, which leads to inflated estimates of switching. Therefore, elasticity approach yields a measure of switching where category sales are held constant (Van Heerde et al., 2003). The intuition is that promotion draws consumers to the category, where after consideration of the items available, they may buy a non-promoted item. This increase in purchase incidence is ignored in elasticity decomposition, but not in unit sales approach that accommodates both promoted and non-promoted items within a category (Van Heerde et al., 2003; Van Heerde & Neslin, 2017). The elasticity approach describes the consumer decision process in detail but the unit sales approach is required to assess net losses in cross-brand sales and net growth in category sales (Van Heerde et al., 2004).

Smaller effect directed to switching may imply that promotions are more appealing to retailers than was previously thought (Van Heerde et al., 2003). This means that promotions do not just redistribute purchases within category, but instead can make categories grow temporarily (Van Heerde, 2005). While unit sales approach shows the net effect correctly accounting for increased category volume, Van Heerde et al. (2003) suggest that one should view the decomposition approaches as complementary measures of sales promotions effectiveness. Elasticity decomposition works at its best for separate assessment of changes in the consumer decision probabilities, ceteris paribus, and therefore it can help to determine for instance, how much the expected number of purchase incidents changes during a promotion for an item (Van Heerde et al., 2003). The advantage of unit sales approach is that it provides a very practical approach underlining the importance of understanding and rigorously analysing promotions effects by means of unit sales effects that managers are accustomed.

2.3 Positioning of the study within the discipline

Over the last decades, sales promotions have appeared in numerous academic papers ranging from marketing, through economics, to psychology (Blattberg & Briesch, 2012). Therefore, it is important to explicate and point out how this study positions itself in the context of sales promotion literature. Sales promotions have been studied from many angles depending on the context. Figure 3 puts together what has been discussed throughout this chapter and illustrates how the research problem of this thesis is positioned within the sales promotion literature.

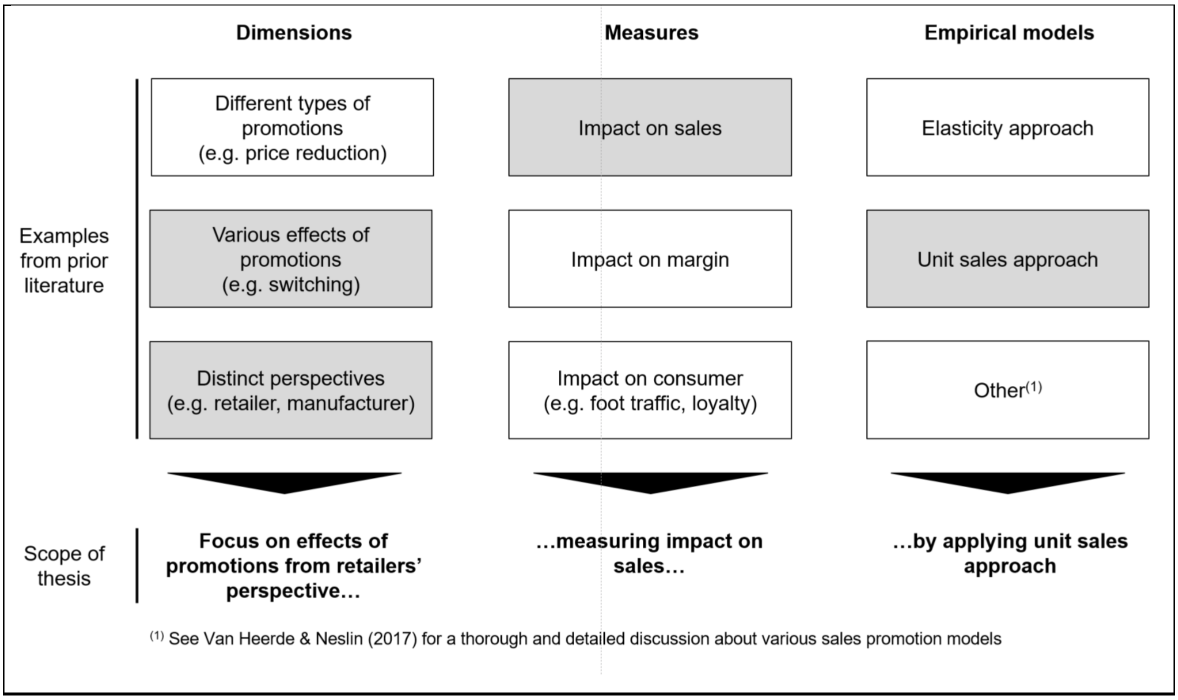

Figure 3. Positioning of the thesis in the context of sales promotion research

First, this study does not try distinguishing between different forms of promotions. This study untangles the effects of promotions from a retailer's perspective regardless of promotion type. The impact is measured on sales, which leaves profit considerations as well as consumer behaviour effects out of the scope. Lastly, the unit sales approach is applied to model promotion effects.

2.4 Summary of key concepts

Often the terminology associated with promotions might remain vague and vary to some extent across studies depending on the context. To reckon with any confusion with the terminology, Table 5 goes through the key concepts necessary to understand the context of this paper. These definitions are compiled for this thesis and intended to clarify the core concepts for the reader.

Table 5: Summary of key concepts

Item	Description
Baseline sales	An estimate of unit sales of an item if no promotion was in place
Gross lift	Unit sales of an item during a promotional period minus its baseline sales for that period
Switching	Portion of gross lift that consumers switch away from other non-promoted items in the same category
Stockpiling	Portion of gross lift that is borrowed from other time periods due to consumers stocking extra quantity for future use
Short-term effect	Effects occuring during the promotional period
Long-term effect	Effects occuring after the promotional period
Primary demand	Timing acceleration and quantity effects of promotions
Secondary demand	Cross-brand effects i.e. switching
Purchase acceleration	Comprises timing acceleration (buying early) and quantity acceleration (buying more). Results in faster consumption or stockpiling

3 Methods and data

3.1 Overview of methodology

Evaluation of the effects and success of promotions have turned out to be continuous challenge and a complex task for retailers (Ailawadi & Gupta, 2014; Dipl-Kfm, 2015). One of the conceivable reasons is the lack of analytical capabilities or organizational practices to manage the growing amount of data. Data driven decision making (DDD) is the practice of resting decisions on the analysis of data, rather than purely on intuition (Provost & Fawcett, 2013). For too long managers have relied on their intuition to make decisions as evidenced by research on analytics, which suggest that 40 percent of major decisions are based not on facts, but on the manager's gut feeling (Davenport, Harris, & Morison, 2010).

For example, retailers could justify their promotional decisions based on experience and gut feeling. On the contrary, the selection of products to promote could be solely based on data analysis of how effective different promotions are. While managers tend to use combination of these, it is quite evident that nowadays companies increasingly are driven by data analytics (Barton & Court, 2012; Provost & Fawcett, 2013).

Emphasizing decision making based on data and business analytics allow companies to make better, faster decisions in their day-to-day business and can significantly improve productivity (Brynjolfsson, Hitt, & Kim, 2011; McKinsey&Company, 2013a). Because of vast amount of data, managers can measure, and thus know, radically more about their businesses and translate that knowledge into improved decision making and performance (McAfee & Brynjolfsson, 2012).

However, measuring effectiveness of promotions is not easy, as retailers' need to manage vast amount of data through item-level promotions in hundreds of categories carrying out as much as hundreds of thousands of promotions each year. Moreover, many challenges and risks are involved in data-oriented projects. For example, data quality can be huge problem that often, however, can be mitigated to some extent with rigorous data preparation. Another pitfall is that companies may forget to define clear business goals that each data mining task tries to solve, leading to "big data, little brain" phenomenon (Gaudiano, 2017). On the other hand, managers should also brace themselves for situations where the data gives unpredictable results or does not give evidence. Even more challenging, yet one of the most important tasks, is to translate the analysis into tangible actions and act upon evidence to deliver improved performance (Boston Consulting Group, 2014).

Often well-understood process that places a structure on the problem, allows reasonable consistency, repeatability, and objectiveness in the outcomes for data-oriented projects (Provost & Fawcett, 2013). Cross Industry Standard Process for Data Mining (CRISP-DM) describes commonly used tasks that many data scientists use to tackle problems (Shearer, 2000). The CRISP-DM framework is illustrated in Figure 4. Although not explicitly presented in the text, this iterative process model serves as a conceptual tool guiding the data analysis process throughout this thesis.

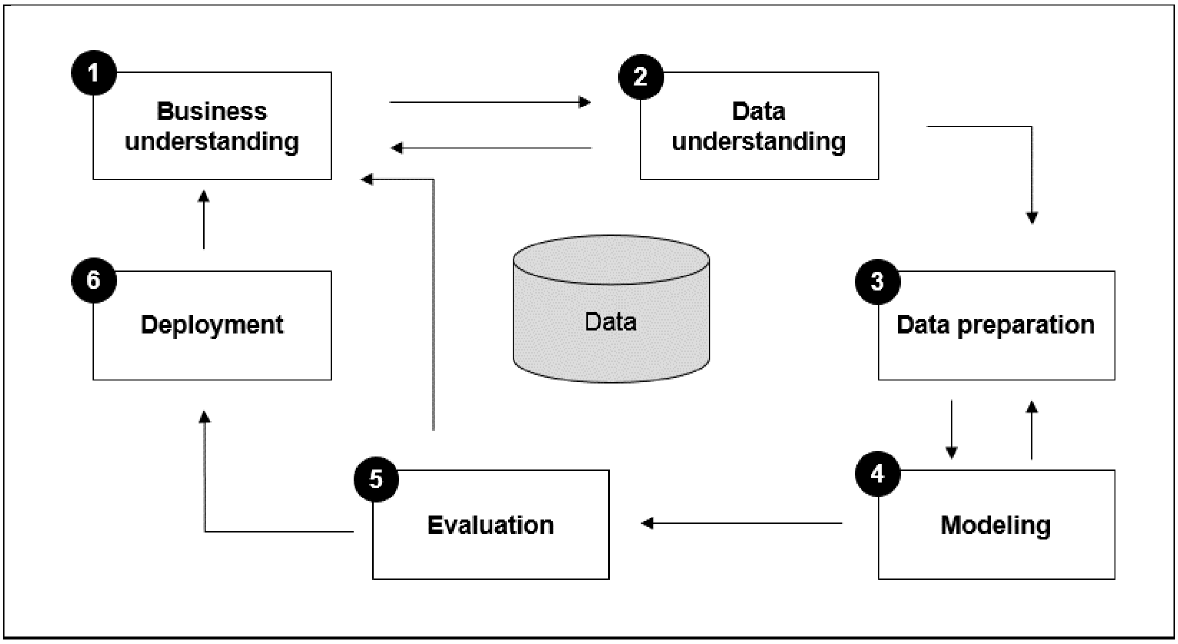

Figure 4. CRISP-DM process model (Adapted from Shearer, 2000)

Understanding the performance of promotions is not an easy task, since the necessary analysis is complex and because often processes are interdependent (Boston Consulting Group, 2014). To measure the impact of a sales promotion, one first needs to evaluate what the sales would be without the promotion. In other words, retailers need an estimate of the baseline sales to understand the sales lift from a promotion. Once the sales in the absence of promotion are estimated, retailers can compare an item's unit sales during the promotional week to its baseline for that week to quantify the gross lift. Thus, gross lift describes the amount of additional sales that the promotion generated.

However, measuring the gross lift does not suffice alone for a retailer, because not all the additional sales is necessarily beneficial from a retailer's perspective. Increase in sales of a promoted item could from various sources that may not all be advantageous. Like discussed in section 2.1.2, the sales lift could come from e.g. increased consumption, other items losing sales within the category or selling less in other time periods (Ailawadi et al., 2006; Van Heerde et al., 2004). If a promotion induces new users to buy a product or existing users to consume at a faster rate, this results in increased total category consumption that is beneficial for a retailer as it creates greater sales volumes altogether.

On the other hand, the sales increase may also be due to switching from other non-promoted brands within the category. For example, if a coffee brand is on promotion, consumers may switch from the other non-promoted coffee brands. Even though the promoted item exhibited a sales lift, the additional sales could be merely due to switching from non-promoted brands. This would not be beneficial for a retailer if the total coffee sales did not increase (except for possible differences in margins). To illustrate, let us assume a store has 10 coffee brands of which Brand A and Brand B are on promotion in a given week. Brand A and Brand B have baseline sales of 10 and 20 units in that week whereas the promotional sales for these items are 15 and 30 units, respectively. Hence, the gross lift of the promotions are 5 units for Brand A and 10 units for Brand B. Now, if the total units sold in the coffee category increases by 15 during the promotional week, which is the sum of the gross lifts of Brand A and B, then no consumption is switched from any of the non-promoted brands. Hence, the switching percentage is zero. Alternatively, if the total category sales would increase only by 10, then 5 units are switched from other non-promoted items resulting in switching percentage of 33.33%. In this paper, short-term incremental lift for a retailer is defined as the gross lift of a promotion after accounting for switching effects. Short-term incremental lift presents with a more realistic measure of promotion effectiveness in comparison to measuring sales lift, let alone investigating only gross sales during promotional week.

Furthermore, promotions can also induce consumers to buy extra inventory for future use. Like switching, the impact of stockpiling is likely not beneficial for a retailer, because it often results in decreased volumes after the promotional period. In addition, promotional margins are commonly significantly lower than with normal prices, which amplifies the negative effect of stockpiling for retailers. Therefore, if customers stack up their coffee reserves during the promotional week, this may decrease sales in subsequent weeks, which means that future period sales may be cannibalized via stockpiling. One should note, however, that the product should be “stockpileable” for this effect to occur. Hence, stockpiling effects may not be present in all product categories. For example, one is unlikely to buy extra quantity of perishable products or clothes despite significant promotion.

To measure promotion effectiveness and to untangle the portion of the additional sales that is truly incremental for a retailer, one needs to account for these switching and possible stockpiling effects. Before proceeding to data description and specific methods that are used to model the promotion effects, Figure 5 presents an overview of the methodology defining each step in the data analysis process.

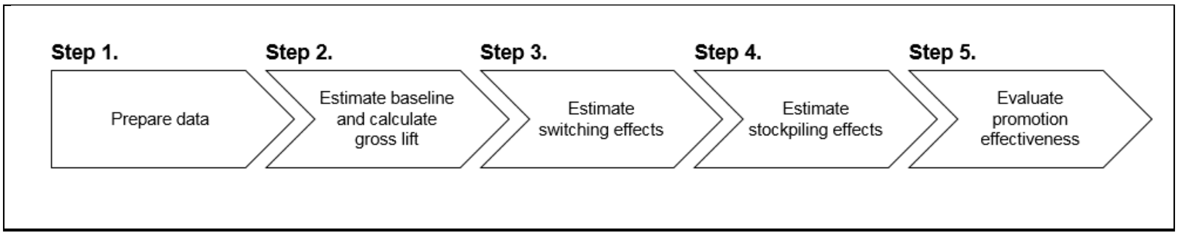

Figure 5. Overview of research methodology

3.2 Data

Generally, there are three types of sales data available for modelling promotions: retailers’ POS data, panel data from a sample of consumers provided by market research companies and customer purchase data (i.e. loyalty card data) (Blattberg and Briesch, 2012). Many studies have focused on store-level models for sales promotions effects, as it tends to be more readily available, more representative and easier to process in comparison to household data (Van Heerde et al., 2004). For this study, POS data provides the needed sales information for addressing the research problem while being simple yet managerially relevant alternative to model promotional effects on sales. Customer purchase data and panel data do not serve similarly for this purpose, as individual purchase behaviour is beyond the scope of this thesis.

In cooperation with a Finnish retailing company, a dataset consisting of 2-year weekly POS data (2015 and 2016) was obtained. A granular data like this is required to answer the research questions and conducting the planned analysis, as it tracks promotions on item as well as product category level. The company manages promotional campaigns centrally, which means that homogenous sales promotions are implemented across its stores in any given week. Therefore, in this study, the store-level differences have only minor practical significance, which allows the use of aggregated POS data across all stores. Moreover, Ailawadi et al. (2006) found out that cross-store variation accounts only about 2% of total variation in net unit impact of a sales promotion for a retailer, making it reasonable to assume that pooled POS data is sufficient to provide reliable results. Blattberg and Briesch (2012) who argue that the decision about the aggregation level of data seems to be driven by whether the aggregation level can answer the research question, and not by biases introduced by the level of aggregation, further support this.

The database where the raw data was located organizes all products in a hierarchy of categories, subcategories and individual items which are consistent across all stores. As there are multiple levels of product categories to choose from, one should decompose the promotional effects at a suitable product category level. Blattberg and Briesch (2012) define category as “a set of products which consumers perceive to be close consumption substitutes”. Using this as a guiding principle, the most suitable level for the analysis was selected among five different product category levels maintained by the company.

3.3 Methods for modelling promotion effects

3.3.1 Baseline unit sales and gross lift

Promotion management begins with the definition of clear, promotion-specific goals (Dipl-Kfm, 2015). Consequently, one needs to be able to track promotion effects at an appropriate level. A basic building block for evaluation of promotion effectiveness is the sales lift from individual promotions; a sales volume for the item on promotion that is attributable to the promotion, which occurred because of the promotion and would not have occurred if the promotion had not been run (Abraham & Lodish, 1993). Therefore, a retailer cannot evaluate the effects of their promotions reliably until baseline sales are computed. Once the baselines are estimated, a gross lift from a promotion can be quantified by comparing actual unit sales to its baseline for that week. This comprises the first research question defined in Chapter 1:

1. **Do promotions generate a lift in sales during the promotional period?**

The bedrock of the empirical analysis in this thesis is a baseline algorithm, which objective is to project what sales would have been during promotion-affected weeks had those promotions not been run. There are different strategies to go about estimating baselines. For example, one could use sales in the previous week prior to promotion, 52-week averages or last year seasonal average (all eliminating promotional weeks) to calculate baselines (CrossCap, 2017). Like Abraham & Lodish (1993) and Ailawadi et al. (2006), this paper estimates baseline sales as a moving average in neighbouring non-promotional weeks. The algorithm projects from “normal” weeks (i.e. weeks that had no promotions), implicitly assuming all other elements as well as competitor moves ceteris paribus.

Naturally, the objectives pose certain requirements for modelling. Since the purpose is to model the effects of tens of thousands of different items running promotions concurrently, the method must be able to perform on a large scale. This also means that manual adjustments are not feasible in this situation. R programming language has gained substantial popularity in analytics in the last decade (Davenport, 2017) driven by its performance capabilities and high-level of flexibility. Therefore, R will be used to develop the baseline algorithm as well as to conduct all other data analysis steps needed in the remaining research process. R as an open source tool provides vast number of packages and ready-made functions to modellers. However, it was necessary to develop a custom algorithm that meets the criteria of this research problem. Simply put, the code runs through the transactional data, and if an item has been on promotion, a custom baseline function is called which then calculates an estimate for that week based on neighbouring non-promotional weeks.

For the baseline algorithm to function properly, one needs to determine suitable number of lagging and possibly leading weeks to be included in the moving average calculations. Two approaches will be experimented to determine the relevant number of lags and leads. First, a fixed number of lagging weeks will be used to estimate baselines. This is likely to produce reliable estimates, as sales prior to promotion are yet to be affected by the promotional activity. However, in some cases, fixed number of lagging weeks may not provide robust estimates due to different characteristics of products. Like Ailawadi et al. (2006), an alternative baseline method was developed that determines the number of leading and lagging weeks based on seasonality and turnover to account for differences in products. Naturally, this second method increases the level of complexity to some degree. To determine which approach to follow one needs somehow to check the robustness of each method. In the empirical analysis, both baseline methods will be tested against data from non-promotional weeks. The one that provides more accurate forecast on the actual data is determined as the more robust alternative and will then be used in the remaining analysis. After producing baseline estimates, the gross lift of a promotion is calculated by subtracting the baseline estimate from the actual unit sales in the promotion week.

3.3.2 Switching

To quantify whether promotions cannibalize sales of other non-promoted products within a category, one needs to estimate the extent to which consumption is switched away from competing products. The second research question addressed this problem:

2. To what extent is the sales lift switched from other non-promoted products within a category?

If for every unit increase in the gross lift of all promoted items in category c in week t, there is a corresponding unit increase in total category units in that week, then, the promotion is not switching any sales from non-promoted items in the category. On the other hand, if the gross lift is purely due to switching from other items in the category, there should be no increase in total category units. Hence, one minus the estimated slope from a regression of weekly category unit sales on the weekly category gross lift is the percentage of the gross lift that is switched from other items in the category (Ailawadi et al., 2006). I will apply a similar approach to Ailawadi et al. (2006) in modelling switching effects. However, I will depart slightly from their model in that the random effects model is replaced with a linear regression model. Since differences among stores were beyond the scope of this thesis, hierarchical modelling such as random effects regression is not necessary and can be replaced with linear regression.

The specific model for switching effects is as follows:

$$Category\ unit\ sales_{c,t} = \beta_0 + \beta_{1c}\text{Total gross lift}_{c,t} + \varepsilon_{c,t} \tag{1}$$

This linear regression model is estimated separately for each product category. The dependent variable is unit sales in category c in week t, the independent variable is the sum of gross lifts across all promoted items in category c in week t, β_0 is the intercept term, β_{1c} is the regression coefficient for category c and $\varepsilon_{c,t}$ is the random error term. Therefore, the switching percentage for a category is given by one minus the coefficient estimate from the fitted regression model:

$$Switching\%_c = 1 - \beta_{1c} \tag{2}$$

where β_{1c} is the coefficient estimate for category c obtained from equation 1.

3.3.3 Stockpiling

The third research question focuses on the effects of promotions after the promotional period:

3. **Do promotions affect future sales via stockpiling?**

To study whether promotions affect future sales, similar approach to switching is applied. Stockpiling effects are estimated from POS data for each category using linear regression. Post-promotion category sales are regressed on the total gross lift across all promoted items in category c. Time window for post-promotion effects is selected large enough so that it includes long-term effects of promotion. The stockpiling model in this paper uses 6 leading weeks, which is supported by earlier literature. For example, Van Heerde et al. (2000) found that 6 weeks suffices for majority of products. Similarly, another study suggests that similar stockpiling estimates can be obtained whether they are calculated based on 4, 5 or 6 periods (Macé & Neslin, 2004). In addition, Van Heerde et al. (2004) used 6-week time window in their model to capture dynamic effects of promotion.

The specific model for estimating stockpiling effects with POS data is as follows:

$$Total\ gross\ lift_{c,t} = \beta_0 + \beta_{1c}\text{Category unit sales}_{c,t+1} + \ldots + \beta_{1c}\text{Category unit sales}_{c,t+6} + \varepsilon_{ct} \quad (3)$$

Like switching model, this linear regression is estimated for each product category. The dependent variable is the sum of gross lifts across all promoted items in category c in week t, the independent variables are category sales in week t+1 through t+6, β_0 is the intercept term, β_{1c} is the regression coefficient for category c and $\varepsilon_{c,t}$ is the random error term. If stockpiling effect were present, one would expect to obtain statistically significant relationship between gross lift and category sales in the weeks after the promotion. On the other hand, if the model does not indicate any relationship between total gross lift and category sales in subsequent weeks, then it is likely that there are no stockpiling effects present.

In conclusion, the promotion effects studied in this paper are modelled at category level, that is consistent with Van Heerde et al. (2004) and Ailawadi et al. (2006). One should note however that even though the switching and stockpiling effects can only be estimated for a category, the baseline algorithm presented in 3.3.1 enables calculation of gross lift on item level, thus accounting for individual promotions. This means that with the help of category-level estimates (e.g. switching), one can evaluate the effectiveness of individual promotions in that category.

4 Empirical analysis and discussion

4.1 Data preparation

When dealing with large volumes of data, considerable amount of time is spent on ensuring quality of data and preparing data for analysis. This thesis was no exception. Data preparation calls for in-depth understanding of the available information early in the process involving both business understanding and common sense to a large extent. Consequently, several tasks were identified and performed to prepare the data for subsequent modelling. Figure 6 lays out the main data preparation steps:

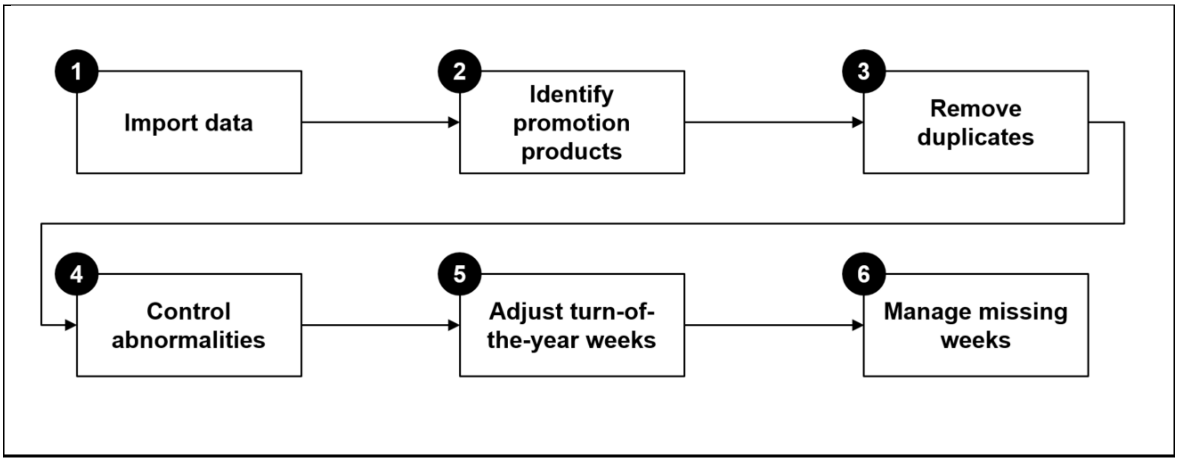

Figure 6. Overview of data preparation steps

Raw sales data were provided in separate text files for each year (2015 and 2016). The data was obtained from a transactional database maintained by the retailer company. As a first step, data was imported to R and the different years where combined into a single data set. At this point, it was noticed that file encoding mixed up the product codes, so one needed to reset the encoding of the files to ensure seamless import. Raw data comprised weekly transactions covering all products, so the next step was to identify items that had been on promotion during the study period. Only items that had been on promotion at least once during the study period was retained in the data set.

During the process, it was noticed that if a promotion had been rolled out in two separate catalogues there was duplicating observations in the data for those weeks due to different campaign names. The corresponding observations were otherwise identical except for the campaign name. Hence, any duplicated campaign sales data were identified and removed from the data set at this point. When further analysing the data, one could also observe that there were negative unit sales recorded for some items. After discussions with the company, it turned out that these abnormal negative unit sales refer to weeks in which no items had been sold but the items were either returned or replaced through warranty. Even though these observations represent less than 0.1 percent of total observations and do not likely to have significant impact on results, these were handled by imputing zero sales for abnormal weeks in question to preserve quality of analysis.

During the initial data exploration, it was noticed too that the last week of 2015 were recorded both in 2016 and 2015 datasets. The most likely reason for this is that these weeks represent only partial weeks at turn of the year. When inspecting the data more in detail, this indeed was the case. As the data from both years were examined, it became obvious that 2015 data included sales from Monday to Thursday whereas the sales for the week 53/2015 in 2016 data were from Friday to Sunday. These partial turn- of-the-year weeks were merged together for each item to represent the actual situation correctly.

Lastly, if a certain item had not been sold in a given week, there was no observation for that week in the data. So, for example an item may have had sales in week 05/2016 but the next observation in the data was not 06/2016 but 07/2016 instead, if the item had no sales in week 06/2016. Baseline algorithm requires unbroken data series to function as intended because it calculates the estimates based on neighbouring weeks. Therefore, one needed to write another R script to generate these missing weeks (i.e. weeks with no units sold) and imputed zero values to these observations so that unbroken time series were ensured for each item.

The pruned data set consisted of over 4 million observations including more than 40,000 different products in total. During the study period of two years, over 140,000 promotions where run in nearly 500 different product categories.

4.2 Promotion effects

4.2.1 Baseline unit sales and gross lift

The starting point for measuring the impact of promotions is to evaluate what the sales would have been without the promotion. Thus, one needs an estimate of the baseline sales. Once the baselines are estimated for each item in each promotional week, one can obtain the gross lift of a promotion as illustrated in Figure 7. Gross lift quantifies the number of additional units that a promotion generates in relation to its baseline sales for that week.

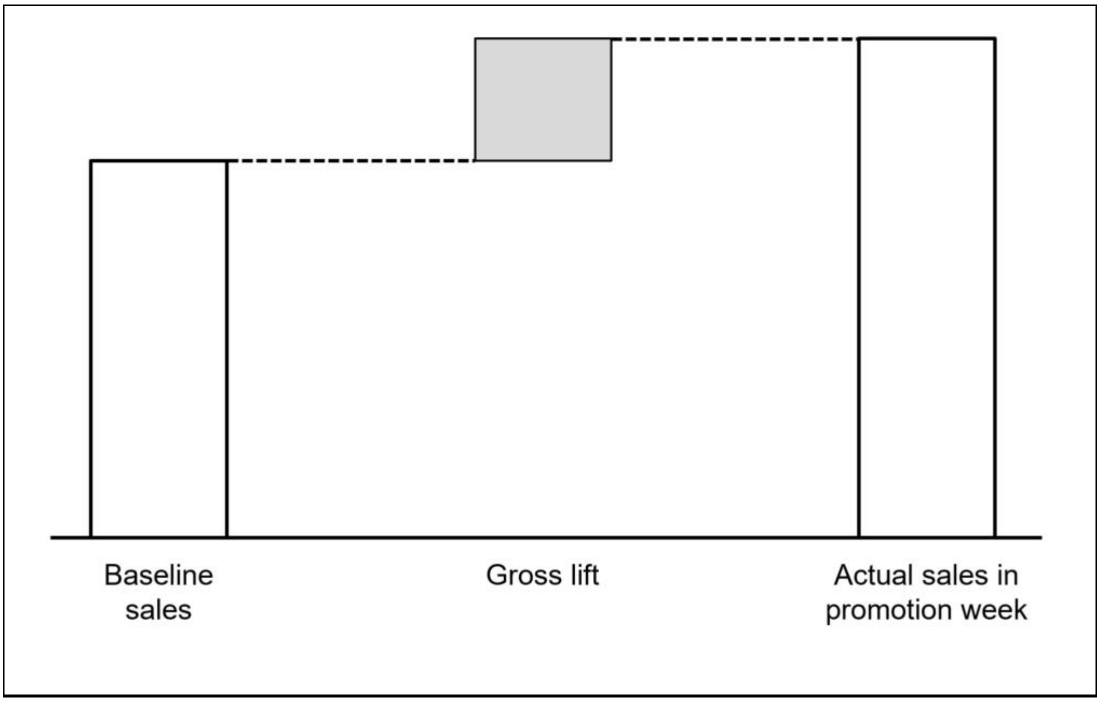

Figure 7. Gross lift of a promotion

Like discussed in the methodology section, the number of lags and leads are needed as inputs to the baseline algorithm which estimates baseline sales as a moving average in neighbouring non-promotional weeks. Two approaches are experimented to determine the suitable number of lags and leads to be used in modelling the baselines. First, a fixed number of lags is used. This means that an item's baseline sales for a promotion week is estimated based on the weeks prior to promotion. The benefit of this method is that the data points that are used in estimation are not yet affected by the promotional activity. Two variations of fixed lags were used, namely 4- and 6-week lags, to estimate baselines. The robustness of estimates produced by these models are discussed collectively once all different baseline models are estimated.

The second approach dictates the number of lags and leads based on the turnover and seasonality of a product. The benefit of this approach is that it accounts for differences in products by determining the suitable number of lagging and leading weeks based on which type of product is in question. For example, longest lag should be used to slow-moving products with a low level of seasonality. The reason for this is that slow-moving items may require more weeks to get an accurate estimate of the sales whereas low seasonality suggests that previous weeks provide the best estimate because these are yet to be affected by the promotion. If an item's sales were highly seasonal, one would likely need to add leading weeks to represent the season for which the baselines are estimated. However, the most representative weeks in general are the ones before the promotion week, which means that one should use leading weeks only when there is high seasonality in item sales.

To determine the suitable number of weeks for the baseline model with dynamic lead and lag values, one needs to categorize the tens of thousands of different items into groups. The grouping is done based on turnover and seasonality. Like Ailawadi et al. (2006) this thesis uses average weekly item sales during the study period to measure turnover. Similarly, coefficient of variation (CV) in monthly unit sales was used to quantify the level of seasonality. CV for each item is given by the standard deviation of monthly unit sales divided by its mean. Since the initial data set consisted of weekly data, a separate data set of monthly item sales was obtained from the company's database. This three-year data set was used to calculate seasonality measures for each item. Like with the fixed lag baseline models, few different models were examined. Three different categorizations for determining the number of lags and leads included in the baseline calculations are put together in Figure 8. In the first model, products are categorized into 2 x 2 matrix based on the median turnover and seasonality in the sample. The second model follows the same logic, but uses slightly shorter periods. The third model is the most complex with 9 different groups. In this model, the high-medium-low grouping was determined based on 25th and 75th percentiles of turnover and seasonality measures.

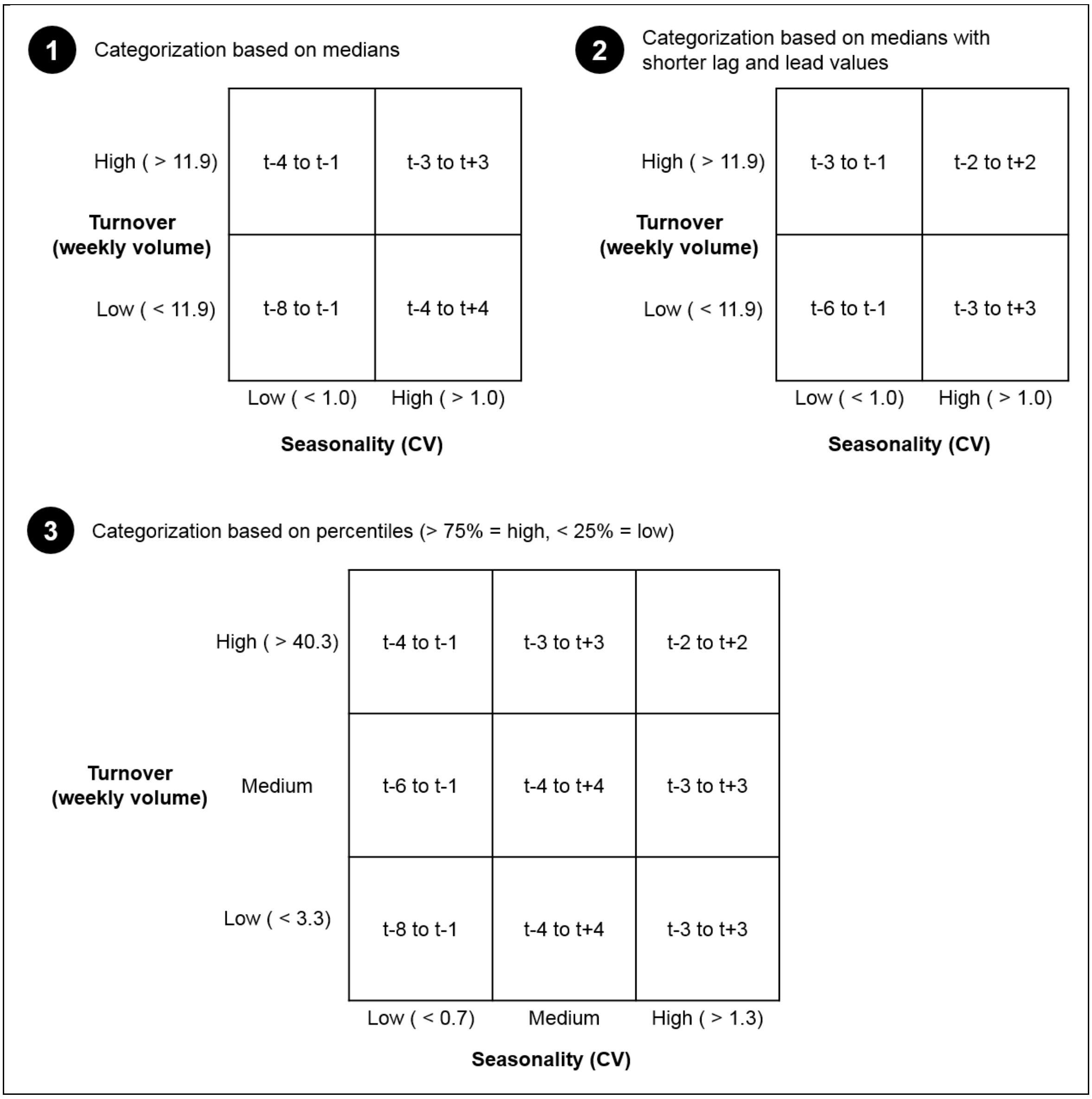

Figure 8. Number of weeks for baseline models with dynamic lag and lead values

Next, one needs to determine which baseline method to use in the remaining analysis. The model selection should be done based on objective evaluation. This thesis used similar approach to Abraham and Lodish (1993) and Ailawadi et al. (2006) by estimating baselines for non-promotional weeks, which allows comparing baseline estimates to actual unit sales. If estimates are close to actuals in non-promotional weeks, one may reasonably assume that the model produces reliable estimates about what the sales would have been if no promotion was in place. The model evaluation encompassed two steps that was repeated for each of the five baseline model candidates. First, a validation data set was extracted and baseline estimates were calculated for non-promotional weeks. At this point, weeks for which baseline sales could not be calculated were removed from the validation set. Moving average is used to calculate the baselines so the number of lags and leads naturally affect how many observations are lost in the beginning and in the end of time series. Hence, baselines cannot be estimated to all weeks and the total number of categories may vary among the models. Second, a paired t-test was conducted to investigate whether the estimates were statistically different to actuals. The null hypothesis is that the true difference in means equals to zero. This indicates that if one fails to reject the null hypothesis, one may conclude that the estimates are not statistically different to actuals and the model seems robust. Like mentioned, the same validation process was performed for each of the five alternative baseline models. The robustness of different baseline estimates measured by the paired sample t-test are reported in Table 6 below.

Table 6: Robustness of baseline estimates

Baseline model	**Robustness p-value**
Fixed number of lags	
4 weeks	0.9043
6 weeks	0.4792
Dynamic lag and lead values	
Categorization based on median turnover and seasonality	0.0007
Categorization based on median turnover and seasonality with shorter lag and lead values	0.0029
Categorization based on percentiles of turnover and seasonality (> 75 % = high, < 25 % = low)	0.0000

Both fixed number of lags models fail to reject the t-test, which means that these two alternatives seem to provide robust estimates. In contrast, all the dynamic baseline models reject the null of zero difference in means having very low p-values. This indicates that the robustness of the dynamic models remains questionable. In this case, adding more complexity by allowing the number of lagging and leading weeks to change depending on the product did not help in producing better estimates. Going forward, it would safe to proceed with either one of the fixed models. However, shorter time periods seem to perform better as indicated by 4-week lag having higher p-value in comparison to 6-week lags. This is further supported by dynamic models where the highest p-value given by the model with shortest lag and lead values. Given the most of the company's products are relatively fast-moving goods, one would expect that longer time periods are not needed for producing accurate estimates. For the subsequent analysis, fixed model with 4 lags will be used. The benefit of using the past month's sales as a proxy for baselines is that it is very intuitive and easy to for managers to understand. In contrast, the downside is that this simple method may not provide accurate estimates for highly seasonal products, which we are likely to have among the tens of thousands of items. Nevertheless, simple models tend to work surprisingly well when dealing with vast amount of data as sometimes adding more complexity only results in overfitting too much "noise" into the model.

Now that the data has been prepared for analysis and the baseline model has been selected, one may delve deep into the first research question:

1. Do promotions generate a lift in sales during the promotional period?

As a first step, the baseline algorithm was run on the data with 4 lags. The baseline algorithm goes through the whole set of data identifying promotion weeks for which the baselines are calculated as a moving average based on the item's unit sales in the previous four weeks excluding any neighbouring promotional weeks. Once baselines were estimated, the gross lift of a promotion was given by the difference between the actual units sold and its baseline for that week. Hence, the gross lift gives the amount of additional sales that is generated due to the promotion in relation to its baseline. In this case, measuring gross lift in absolute terms is rather non-meaningful for two reasons. First, products vary significantly in terms of sales volumes making it difficult to compare products that are very different to one another. Moreover, as the data used in this thesis is pooled across stores, reporting sales lift in absolute terms is not very intuitive and might be hard to interpret from a managerial perspective. Therefore, gross lift is reported as a percentage of baseline sales to give a relatively measure that is more comparable across items and easier to interpret.

Prior to analysing the gross lift in detail, one should note that there are few special cases in which gross lift percentage cannot be calculated. Since the baseline algorithm uses 4 lags in estimation, observations are lost in the beginning of the time series. For example, if a product has been on promotion in week 03/2015, one cannot estimate the baseline with this method as there are only two data points from which to predict (02/2015 and 01/2015). In these events that form about three percent of all promotions, the algorithm identifies that there are not enough data available and returns N/A value. Another special case arises if a product has not been sold in the previous 4 weeks, which leads its baseline estimate to 0, and thus will return infinite value for the gross lift percentage. These cases represent less than 2% of all promotions in the data. The third special case concerns products that have been on promotion over 4 weeks straight. These cases are problematic because the baseline algorithm is tuned to exclude all neighbouring promotional weeks when calculating the baselines. This means that if a product is on promotion for 5th week straight, one cannot predict baselines because all 4 lags are promotion weeks as well. As one might guess, this does not happen too often representing less than 0.1% of all promotions, but however needs to be considered prior analysing the results. To sum up, calculating gross lift percentage for special cases laid out above cannot be done, and thus these observations were removed from the dataset as outliers. The pruned data after removing outliers totals to slightly over 140,000 promotions in 473 different categories.

The first empirical finding is that 75% of promotions are generating a lift in sales i.e. having gross lift that is greater than 0. This means that as much as one quarter of promotions do not generate lift in sales at all or have a negative impact on sales. In their study on promotions, Boston Consulting Group (2015) reports that 20-50% of promotions are not able to generate noticeable lift or have negative impact. BCG does not specify how "noticeable lift" is quantified, but if 10% gross lift were used as a threshold, the results suggest that about 73% of promotions generate additional sales. This implies that a little less than 30% of promotions do not generate noticeable lift in sales. This falls within the range reported by BCG. Table 7 Table 1gives descriptive statistics and Figure 9 presents histogram of gross lifts found in this study. For the sake of visualisation, 95th percentile of gross lifts is excluded from the histogram resulting in about 135,000 observations. The median gross lift across all the 140,000 promotions examined is 127% with the mid-50% of promotions falling between 17% and 381%. The measure is interpreted so that for every unit of baseline sales, approximately 1.27 additional units are gained due to promotion.

Table 7: Overview of gross lift

Gross lift	**% of baseline sales**
Median	126.5
Mean	642.2
Std.dev	7,259.8
1st Q	16.7
3rd Q	381.2

n = 140,184

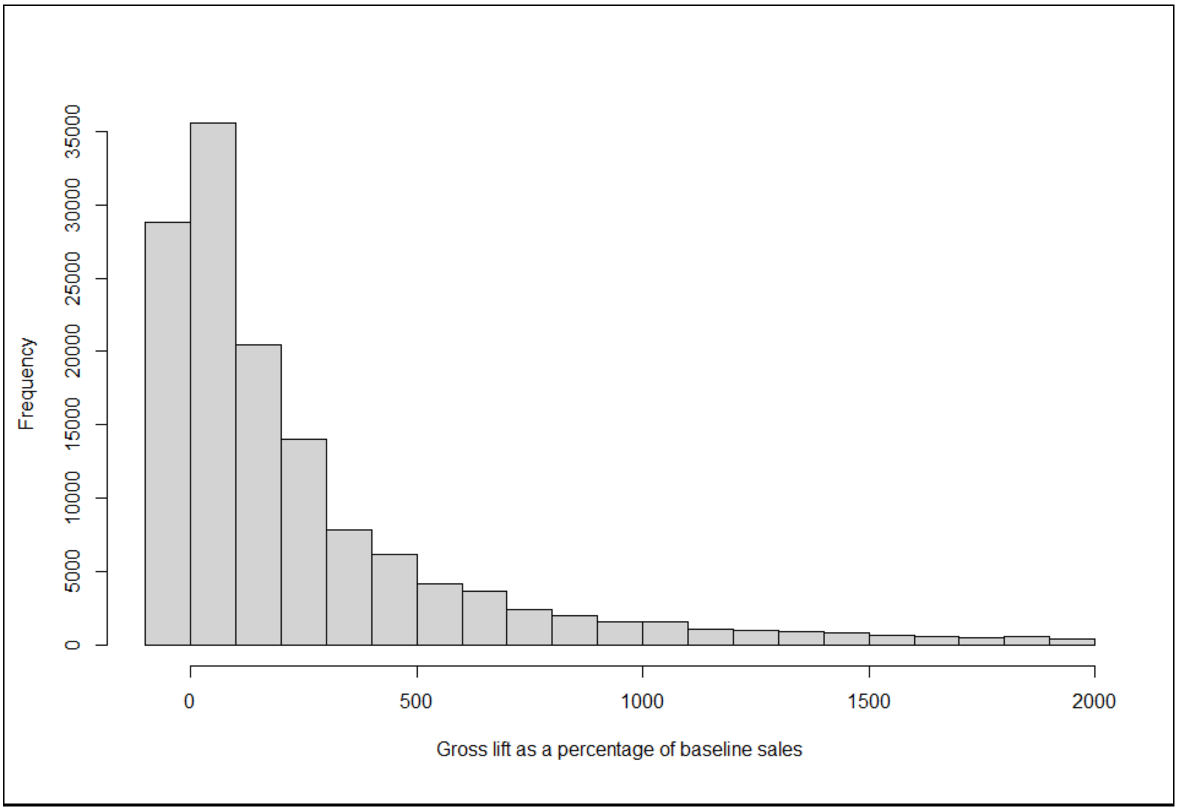

Figure 9. Histogram of gross lift

When grouping the individual promotions based on categories, one can observe that the best performing categories in terms of gross lift are able to generate over 1,000% sales increase which equals to over 10-fold increase in relation to their baselines. Summary of the largest and the smallest gross lifts are presented in Table 8. The largest impact on sales was observed in the fireworks category. However, these products are clearly outliers since retailers in Finland can sell fireworks only few weeks once a year prior to New Year. Estimating baselines based on the prior 4 weeks sales therefore does not reflect the nature of these products, which essentially means that some other method should be used to evaluate the promotion effects in this category. If these types of products were of specifically interest and to be examined, one could derive baselines for example from the same weeks in the previous years to better account for the unique characteristics.

Table 8: The largest and the smallest gross lifts

Top 10 largest gross lifts

Category	Gross lift%
Fireworks	8766
Kids' meals and drinks	2936
Power tools	1329
Paper towels and toiler papers	1141
Cooking vessels and utensils	1141
Breakfast and snacks	1126
Electrical accessories	1050
Flowers	996
Serving bowls	970
Household appliances	880

Bottom 10 smallest gross lifts

Category	Gross lift%
Indoor flower care	-4
Women's accessories	8
Fishing	13
Kids' accessories	14
Men's accessories	22
Childcare	28
Small hadrware	33
Women's socks and tights	39
Alcoholic beverages	46
Women's underwear	52

Note: Highest product category level used for more meaningful interpretation

Since consumer behaviour is affected by numerous interrelated issues such as the type of promotion. depth of discount, and brand and store characteristics, among others, it is hard to untangle exact causal effects about why some products perform better than the others. One would likely need to examine each item much more in detail while controlling for various effects to derive reliable interpretation about the consumer buying behaviour. While the first research question was outlined only to quantify the extent to which promotions can generate sales lift and was answered already, one may devote a moment to reflect why these some of the products might respond to promotions in a way they do. For example, paper towels and toilet papers are typical consumables that are used from one week to another. Considering products like these it makes sense for consumers to react to promotion because they know for sure that they we will need the type of items relatively soon. In contrast, the other categories listed in Table 8 are bit harder to interpret because these are very different in nature. For instance, one typically buys serving bowls significantly less often. One thing that may explain these products to generate much additional sales is that consumer may demonstrate strategic buying behaviour in search of greater savings. Serving bowls can be relatively expensive in comparison to fast-moving goods, which may induce consumers to buy the product when it gets promoted due to greater absolute savings. Another possible explanation may stem from the fact that people want their serving bowls to be of the same series whilst often needing more than one of each kind. Therefore, promotions may give a reason for consumers to buy multiple items at the same time. Correspondingly, substantial gross lift of flowers may be explained simply by impulse buying. Flowers are by no means a necessity and often people buy these items as an extra or maybe to please someone. Thus, due to the nature of the product, promotion may just enough to trigger the consumer to buy. Another interesting result is the behaviour of power tools and electrical accessories because one would think that consumers tend to buy these products only when they need it. Intuitively, one would expect promotions to have only minor impact on sales regarding these kinds of products, however that does not seem to be the case.

From the practical perspective, identifying products and categories that have poor impact on sales may be even more interesting for retailers. If a product does not generate a lift in sales, or has a negative impact, there is no need to further analyse promotion effects like switching. Once the baselines are estimated, one can identify worst performing items in terms of gross lift and take already action without the need for further, often complicated analyses. Identifying poor promotion products early in the process is likely to improve overall performance per se and save time and effort significantly. The empirical findings of this study suggest that accessories such as bags, belts, hats and scarfs are one of the least effective categories in terms of gross lift. For example, women's accessories generate only 8% lift in sales. Similarly, men's and kids' accessories exhibit only 22% and 14% gross lifts, respectively. Furthermore, promoting clothes in general seem to have weak impact on sales as well. There are tens of thousands of promotions implemented in these categories which however are generating only moderate lift in sales. For example, in women's underwear category there has been nearly 15,000 promotions during the study period generating only 50% lift in sales on average. This is significantly less than the median gross lift of 127% of the total sample reported earlier. In comparison, hair care products generate about 250% gross lift on average. One should note however that hair care products are promoted much less. This may suggest that great number of promotions dilutes the impact of a single promotion suggesting that retailers should put effort in identifying the right balance not only in terms of which products to promote but also how often.

Another interesting finding relates to fishing products. These have been heavily promoted but impact on sales, however, has been negligible and in some cases even negative. This suggest that the retailer should consider cutting at least some promotions in these categories straightaway. Among the poor performing products are also small hardware items such as screws, hinges and locks. Due to nature of these products, one would think that the promotion alone is not likely to attract the customer to buy the product if there is no explicit need for it. Another interesting finding is that promoting alcoholic beverages generates only minor impact on sales, as evidenced by relatively small gross lift (46%).

The reported standard deviation of 7,260% in gross lifts indicates significant variation among items. 7,000% gross lift is equivalent to 70-fold increase in unit sales. Although this may sound huge, it is not uncommon for items with relatively low absolute volumes that are likely to exhibit substantial increases in sales when promoted. One product that illustrates the situation well is Iittala Mariskooli glass bowl. It is recognized as a quality brand with a well-known price level. Weekly volumes for a one type of this bowl had been only few units if any when suddenly the sales jumped up to almost 3,000 units during a promotion week. If a baseline was estimated at two units for that week and the corresponding gross lift totalled to 3000, the reported gross lift would be 150,000% in this case. While most of sales lifts found in the sample are not anywhere near as substantial as this example, it illustrates well why large deviations may feature.

What comes to the magnitude of gross lift, the median gross lift of 127% reported in this thesis is in the ballpark in relation to prior literature. According to Gedenk et al. (2010), several hundred percent increases in sales are not unusual. This is supported by Ailawadi et al. (2006) who report 310% median gross lift in their study regarding CVS, one of the leading drug retailers in the US. Furthermore, a study across 164 categories suggest mean gross lift between 34% and 292% depending on the type of promotion (Narasimhan, Neslin, & Sen, 1996).

In this thesis, the findings about gross lift were accompanied with significant variation among items as evidenced by high standard deviation. Ailawadi et al. (2006) report variation to much lesser extent, which might stem from the fact that drug retailers have fewer categories in their assortment. One could reasonably assume that as the number of products and categories increase, the probability of having larger deviations from the average response increases as well explaining the differences in deviation to studies that examine only few products at the same time. Another likely reason for differences in overall variation in comparison to Ailawadi et al. (2006) who also studied all items simultaneously is that they allocated substantial effort to recognize and prune out unusual events from the data. The authors deleted holiday weeks and other weeks in which known unusual events occurred in the market in search for “normal” promotion response. While this may provide more stable results, the process of identifying unusual events involves lots of manual work and collaboration with the many stakeholders from the retailer company yet providing marginal value-add considering the events are not likely to have major impact on the reported results overall. Reflecting on the methodology used in this study, omitting certain weeks from the data may also distort baseline calculations that require unbroken time series. Therefore, pruning extreme events or holiday weeks from the data was left beyond the scope of this study.

In summary, items that are not able to generate lift in sales should be omitted from future campaigns. In contrast, products or categories with high levels of gross lift serve as good candidates for promotions. However, one should recall from the first chapter that measuring the gross lift alone does not suffice. One needs to untangle the sources of this lift to reliably evaluate the effectiveness of promotions because the gained units may not necessarily be truly incremental for the retailer. In the next section, we will switch our focus to quantifying the extent to which the gross lift from a promotion cannibalizes sales of other non-promoted products within a category. This forms an integral part of measuring overall performance of promotions from the retailer’s perspective.

4.2.2 Switching

To quantify whether promotions cannibalize sales of other non-promoted products within a category, one needs to estimate the extent to which consumption is switched away from competing products in that category. Like discussed in the methodology section, if for every unit increase in the gross lift of all promoted items in category c in week t, there is a corresponding unit increase in total category units in that week, then the promotion is not switching any sales from non-promoted items in the category. On the other hand, if the gross lift is purely due to switching from other items in the category, there should be no increase in total category units. Figure 10 gives a simple illustrative example of switching effects in a category having two items on promotion in a given week. The first example illustrates an optimal case where no consumption is switched from non-promoted items. This is a desired situation for retailers because all gained units are incremental and thus promotions are not cannibalizing sales of other non-promoted items. The second example illustrates less effective promotions. In this situation promotions generate 15-unit sales lift, however, the category sales increase only by 10 units in total implying that 5 units are switched from other non-promoted items within the category.

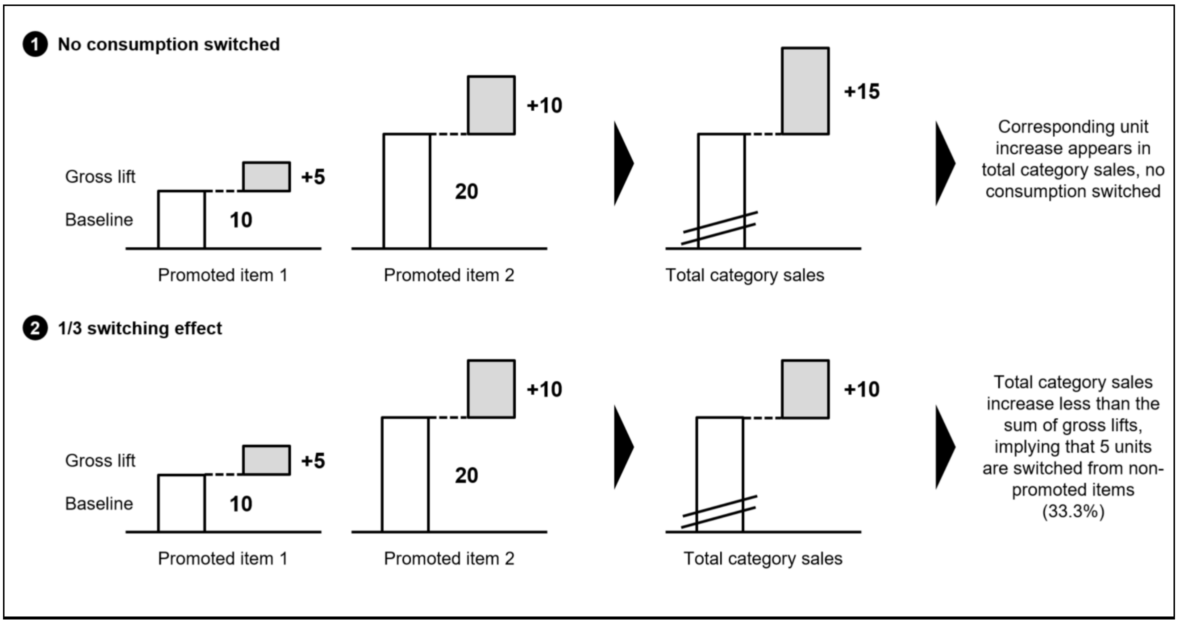

Figure 10. Illustrative example of switching effects

Before fitting the linear regression models for each category, de-seasonalisation of regression data is needed. This stems from the fact that both gross lift and category sales may increase in high season or decrease in off-season resulting in spurious correlation. Hence, de-seasonalisation tackles the spurious effects that otherwise might distort the results. Guided by Ailawadi et al. (2006) de-seasonalisation of the data was conducted using week-of-year seasonal factor. The seasonal factor for category c in week t is the average category sales in week t divided by average weekly category sales. Both the category sales and total gross lift from Equation 1 were divided by the corresponding seasonal factor prior to regression. Additionally, one should note that some categories may not have enough promotion weeks to fit the model reliably. For example, if there are only 5 observations in a given category it does not make sense to fit a regression model into data with such a small sample size. Therefore, categories with less than 30 promotion weeks in total were excluded from the regression analysis. This resulted in 150 different categories for which switching effects can be modelled.

The second research question was defined as follows:

2. To what extent is the sales lift switched from other non-promoted products within a category?

To answer the research question, one needs to fit switching model separately for each of the 150 categories. From Equation 2 we may revise that one minus the estimated slope from a regression of weekly category unit sales on the weekly category gross lift gives us the percentage of the gross lift that is switched from other items in the category.

First, the data needed to be grouped by each category and then linear regression model was run separately for each category. Broom package in R contains three functions that deal with complex returned objects from statistical operations by groups. Broom was utilized to conveniently obtain regression coefficients and other output statistics for each model. Table 9 gives an overview of the reported switching effects and Figure 11 illustrates the distribution.

Table 9: Overview of switching effects

Switching effects	**% of gross lift**
Median	38.50
Max	71.04
Min	0.04
Std.dev	15.80
1st Q	30.30
3rd Q	47.50

n = 139 categories with p-value < 0.01

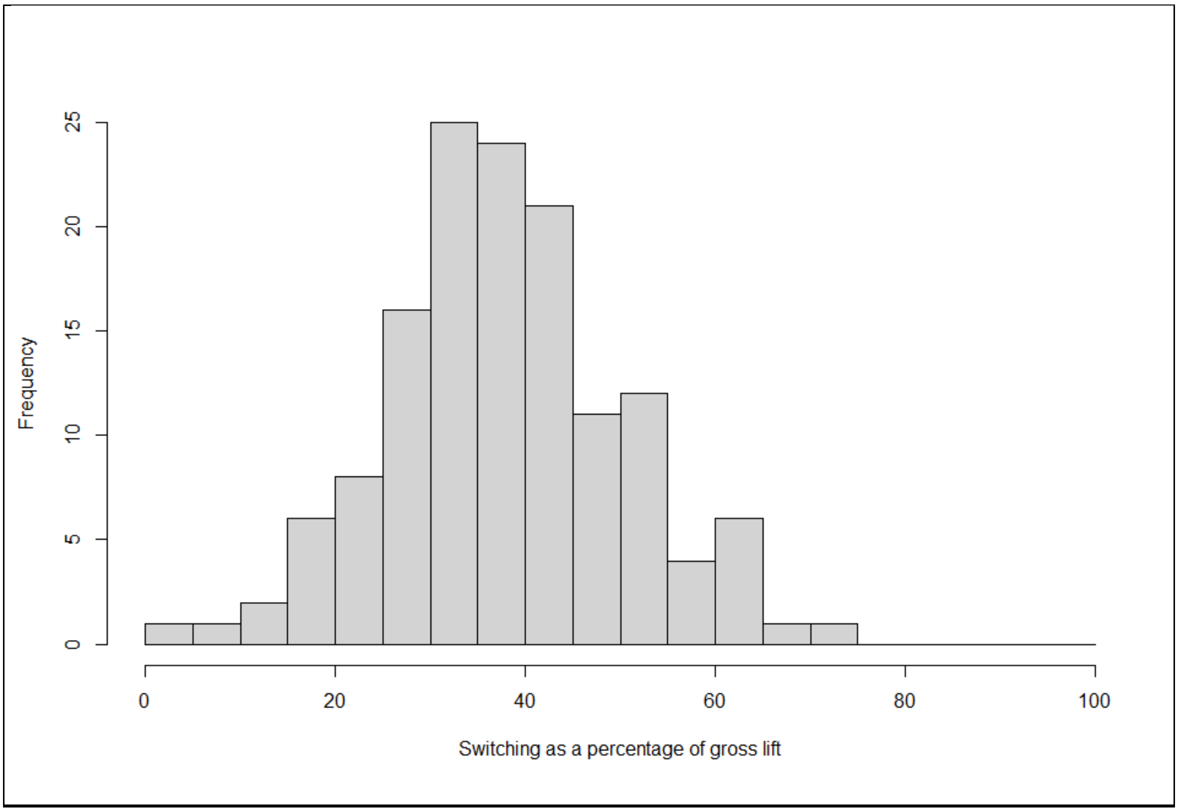

Figure 11. Histogram of switching effects

The empirical findings suggest 38.5% median switching effects. This implies about 60% short-term incremental lift for a retailer. Short-term incremental lift refers to the portion of gross lift that is left after accounting for the switching effects. At 1% percent significance level, 139 out of the 150 models have statistically significant switching effects. If 5% significance level were used, the percentage of categories with significant switching effects would increase to 97%. The results suggest that regression method is functioning as intended and the modelling approach seem robust overall. Further, all the estimated switching percentages fall between 0 and 1. This is a very promising result as the linear regression model in itself does not restrict the coefficient estimates to stay between the practical limits of 0 and 1.

Interestingly, alcoholic beverages are among the few categories that did not have statistically significant switching effects. For example, the coefficient estimate for beers indicates large effect size; however, the model fails to reject the null at 1% and 5% levels with a p-value of approximately 0.06. Considering the buying behaviour regarding beer products a high variation within the sample was hypothesized because people, especially in Finland, tend to buy large amounts of beer from time to time. In fact, a few outliers can distort coefficient estimates from the regression model by inflating the standard deviation significantly. Even with large effect sizes the high variation results in lower t-statistic value which naturally lifts the p-value respectively. If there are significant outliers present, this may cause the researcher to fail to reject the null hypothesis indicating that the coefficient estimate is not statistically significant. To test this hypothesis, the regression input parameters were examined more closely for this certain category to see if there are any outliers or extreme events present in the data.

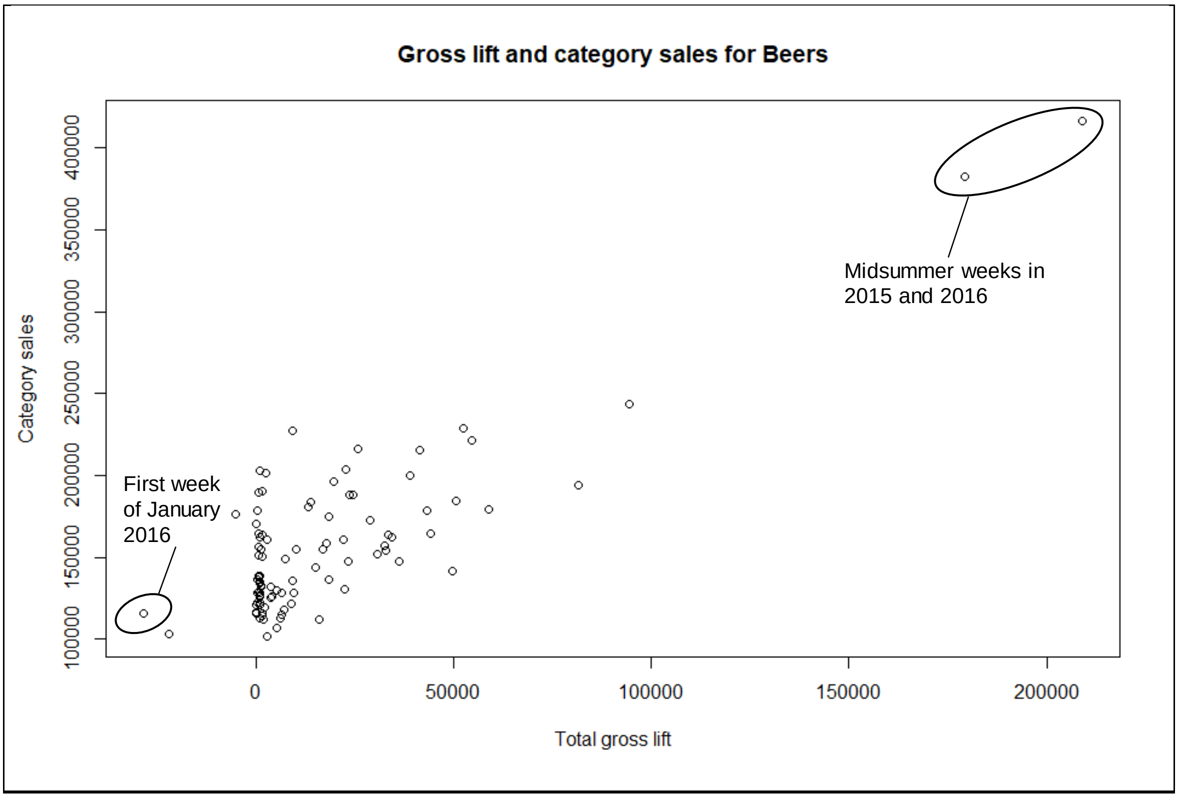

Figure 12. Outlier detection for Beer-category

Figure 1Figure 12 above presents a scatter plot of total gross lift and category sales in beer category. One may see from the plot that there are two observations clearly standing out from the rest. Interestingly, these two observations for which a clear spike in demand and correspondingly in gross lift was observed are the midsummer weeks in 2015 and 2016. It seems that during these weeks people have bought significantly more beer. As a Finn, this is rather easy to comprehend as Midsummer is one the most popular holiday events in Finland. What happens in modelling is that the baseline method, which uses moving average based on previous month's sales, is not able to account for this extreme event resulting in inflated gross lift values, which then again affects the regression model by increasing variation in coefficient estimates. Apart from the two midsummer weeks, two other extreme values also caught one's attention. It seems weird to have large negative gross lifts when otherwise promotions in this category appear to generate additional sales. When examining further, it was noticed that the largest dip in gross lift was observed during the first week of January. The reason for this is that the week follows a period of holiday events during which beer sales are likely to go up significantly, such as the New Year's Eve. This in turn inflates the baseline estimate, as opposed inflated gross lifts during Midsummer weeks, which for its part distorts the results as well. To account for these extreme evens, these three weeks were removed from the sample as outliers. Regarding the other negative observation found in the scatter plot, similar justification could not be formed so it was kept in the sample. After adjusting for the extreme events, the coefficient estimate's p-value dropped from about 0.06 to 0.00 suggesting 57% switching effect for the category. This confirms the hypothesis that few extreme events skewed the initial results.

For being consistent, same investigation was conducted for the other two alcoholic beverage categories, namely ciders and long drinks that may exhibit similar behaviour. The regression data was plotted for both categories in search of outliers. Some potential outlier candidates were found, however, the other observations in general were in both categories much more scattered in comparison to beer category. As one might guess, standard deviations in these cases may already be relatively high due to wide spread of observations so removing outliers is not likely to transform coefficient estimates statistically significant. In fact, this was the case which means that these categories must be excluded from the analysis as one cannot reliably infer switching effects for the coefficient estimates. Similar hypotheses to the behaviour of alcoholic beverages could not easily be formed regarding the few other categories that did not exhibit statistically significant switching effects. Thus, one needed to exclude these categories from the discussion.

Table 10: The largest and the smallest switching effects

Top 10 largest switching effects		**Bottom 10 smallest switching effects**	
Category	***Switching%***	***Category***	***Switching%***
Barbecues	71.04	Thumb drives and memory cards	0.04
Men's long-johns	66.00	Pneumatic machines and devices	7.14
Kids shirts and college shirts (pre-schooler)	64.10	Flours	10.22
Cold cuts (e.g. pork, turkey)	62.66	LED lamps	13.39
Kids outdoor clothing (pre-schooler)	62.88	Men's electric shavers	15.49
Kids socks (school-aged)	60.84	Electric toothbrusts	16.68
Barbecue sausages	60.06	Bathroom tissues	17.72
Beers	57.43	Sweeteners	19.23
Pastries	55.72	Pots and pans	19.82
Windshield washing fluids	55.50	Tinned meats	19.91

The largest and the smallest switching effects are reported in Table 10 above. The largest switching effect was found in barbecues which was slightly over 70%. This means that if an item generates 10 additional units in comparison to its baseline, only about 3 of these represent incremental sales and the rest 7 units are switched within the category. Men’s long-johns and small kids’ shirts exhibit large switching effects as well. Intuitively, this makes much sense because these kinds of products may not be tied to any certain brand. Imagine a mom buying a shirt for a small kid. One can assume that it does not matter too much which brand to buy because small children are often not that interested about their clothes. Thus, a mom buying a shirt for her kid may easily switch between brands if the other shirt is on promotion. Similarly, men’s long-johns are worn underneath jeans or other pants so the brand does not matter much implying that high level of cannibalization may occur. Beers also exhibit high switching effects between brands. This imply that customers may easily buy e.g. Koff instead of Karhu if the former is on promotion. In general, Finnish beer brands exhibit relative small differences in terms of taste (although some may argue differently) so the results are rather easy to comprehend. Overall, a high level of switching effects implies that even though these products may exhibit positive gross lift, the effect of promotion is diluted due to cannibalization caused by brand switching within a category. What retailers would want to see is a low level of switching effects resulting in greater incremental sales.

Smallest switching effects by category are also reported in Table 10. These categories represent good candidates for promotions because most part of the gained units are incremental for the retailer. It is quite difficult though to explain to why exactly these categories are behaving the way they are and this formulates a topic for another study per se. However, by systematically identifying categories that cannibalize sales of other products to a lesser extent allow retailers to direct promotions to categories where they are more effective.

A key takeaway from these findings is that switching effects are largely present in the data and retailers need to consider this when making promotional decisions about which products to promote as well as when evaluating promotion effectiveness. Naturally, these decisions connect with the overall objectives that the retailer has been committed to. If one merely targets to increase store traffic as such, one may neglect switching effects associated with individual promotions. In this case some other method for measuring the effectiveness is though needed. This thesis provides retailers a holistic approach to investigate performance regarding promotions across multiple categories, which oftentimes is of managerial interest.

The analysis above investigated switching effects for 150 product categories but most of the prior literature has concentrated only on few selected categories at a time. Hence, this study provides a great amount of new information about less researched categories. The empirical findings about switching effects reported in this paper are in line with prior literature. This suggests that the results can be interpreted reliably regarding also those categories that may not have been previously studied. Table 11 summarizes the results from this thesis in comparison to earlier literature on selected categories.

Table 11: Comparison of switching results to earlier literature

Category	Switching% This study	Switching% Other(1)	Article
Yogurts	38	37	Pauwels et al. (2002), Van Heerde et al. (2004), Sun
Bathroom tissues	18	21	Van Heerde et al. (2004)
Paper towels	33	14	Chan et al. (2008)
Hair care	35	31	Van Heerde et al. (2004)
Beers	57	15	Leeflang et al. (2008)(2)
Juices	46	8	Nair et al. (2005)
Detergents	32	39	Leeflang et al. (2008)
Fabric softeners	43	29	Leeflang et al. (2008)(3)
Dish detergents	41	28	Leeflang et al. (2008)(4)

(1) Average if multiple studies

(2) Average across bottled and canned beer

(3) Non-concentrated fabric softeners

(4) Non-concentrated dish detergents

Most of the switching effects reported above are in the ballpark with previous studies. Largest difference was found in beer and juice categories. Switching effects found in this study are much greater for these categories than reported by Leeflang et al. (2008) and Nair et al. (2005).

4.2.3 Stockpiling

So far, this thesis has focused on short-term impact of promotions which refer to effects occurring during the promotional period. It was found that most promotions do generate sales lift but on the other hand switching effects often dilute the impact. In addition to short-term impact, retailers may also be interested in whether promotions affect sales in weeks after the promotion. To extend the analysis beyond the promotional period, one may investigate whether the future sales are affected via stockpiling which was the third research question laid out in Chapter 1:

3. **Do promotions affect future sales via stockpiling?**

Promotions can induce consumers to buy extra inventory for future use. Like switching, the impact of stockpiling is likely not beneficial for a retailer because it often results in decreased volumes after the promotional period. Figure 13 gives a simplified illustration of stockpiling effects. To study this effect, linear regression model presented by Equation 3 in the previous chapter is fitted for each category using total gross lift as a dependent variable. Category sales in the 6 subsequent weeks are used as independent variables. The number of leadings weeks were selected based on prior literature. Following similar procedure to switching model, the data was de-seasonalised before running the regressions.

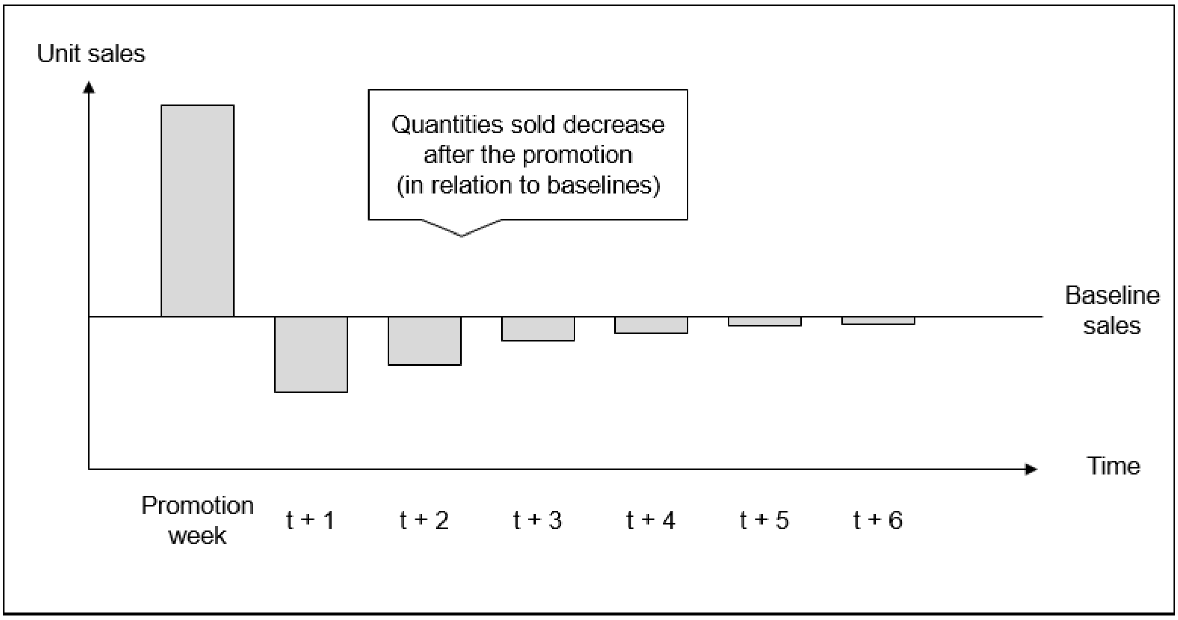

Figure 13. Illustrative example of stockpiling effects

First observation after fitting the models is that only less than 10% of the individual coefficient estimates are statistically significant at 5% significance level. When looking at the results more closely, one finds that there are nearly as much negatively signed coefficient estimates as positive ones resulting in highly inconsistent coefficient estimates. Randomness of the coefficient estimates suggest that there may not be stockpiling effect present. Another thing that may explain random behaviour is the presence of multicollinearity in the applied multiple regression model. Multicollinearity may cause erratic changes in individual coefficient estimates. This happens if the predictor variables, in this case the leading category sales, are correlated with one another. However, this does not reduce the overall reliability of the model. To check whether the overall stockpiling models are significant even though the individual coefficient exhibit inconsistent patterns, one can run F-tests to evaluate whether our regression models fit the data overall. This subsequent investigation suggests similar conclusions to the initial findings. Only 12% of the stockpiling models are statistically significant. When looking at the r-squared values from the various stockpiling models, it seems that the category sales after the promotion period do not explain much variation in the gross lift. All the fitted models explain less than 50% of the total variation. This means that the model either does not fit the data well or there are no stockpiling effects present in the data. The reason may simply be that there is only a weak relationship between the gained units in the promotion week and the sales in the weeks after the promotion or no relationship at all. The main conclusion from the above analysis is that this sample does not give evidence on stockpiling effects and one cannot draw further conclusions about the stockpiling effects based on these results.

Earlier literature suggest that stockpiling effects may not always be present. One of the mysteries of store-level POS data is the lack of dip in quantity sold in the weeks after the promotion often suggested by consumer inventory model (Blattberg, Briesch, & Fox, 1995; Hendel & Nevo, 2003; Van Heerde et al., 2000). Even though many studies relying on household-level data have found stockpiling effects, small effect sizes can be harder to spot in aggregate data (Hendel & Nevo, 2003). This study used a regression model based on store level data which may partly explain the absence of stockpiling effects. In their study, Ailawadi et al. (2006) used the same model in estimating stockpiling effects. They found that stockpiling effects fitted from the store level POS data provided much smaller stockpiling effects in comparison to results estimated from another model using household-level data obtained from a loyalty program. Indeed, their reported median stockpiling effect from POS data was barely above zero.

Moreover, Macé & Neslin (2004) report that post-promotion dips are often related to product characteristics such as high-priced, high-share and frequently promoted products as wells as store area demographics like older customers and large households. Further, all products are not necessarily “stockpileable” from practical perspective. For example, foodstuff items are often perishable and cannot be stored for long periods of time. This naturally prevents customer stockpiling. Another example is clothing products which rarely induce customers to stockpile. Imagine going to a store to buy a sweater. How likely one would buy extra quantity just because the item happens to be on promotion? Therefore, product characteristics are likely to explain at least part of stockpiling effects like Macé & Neslin (2004) argue. Consequently, stockpiling effects may indeed apply only to a certain subcategory of products. Items that are non-perishable and not too expensive may feature stockpiling effects. However, earlier literature and the findings from this study imply that the effects of stockpiling may be overstated in earlier decomposition studies that have focused only on few categories at a time.

4.3 Synthesis of empirical findings

The empirical findings suggest that most of the promotions do generate a lift in sales. In comparison to baseline sales, a median gross lift of about 130% was found. However, a significant part of this promotion impact did come from switching which dilutes the effectiveness of the promotions from a retailer's perspective. About 40% of the sales lift generated by promotions accounts to switching within the category, thus cannibalizing sales of other non-promoted items. Long-term effects of promotions were also studied but no significant stockpiling effects were found in the sample.

Reflecting the results to prior literature, the findings of this paper seem reasonable. However, before concluding the empirical part, a further robustness check is performed. To validate the analysis, the different baseline models introduced in paragraph 4.2.1 were tested to see whether changing the baseline model affects switching and stockpiling results. The initial model used in this thesis estimated baselines based on previous 4 weeks' sales. The alternative fixed model uses prior 6 weeks sales to estimate baselines. In addition to the 6-week lag model, 3 different dynamic models were experimented. In these models, the number of leading and lagging weeks for each item were determined based on product characteristics; seasonality and turnover to be specific. Coefficient of Variation in monthly unit sales was used to quantify seasonality whereas average weekly volumes described turnover. Table 12 gives a summary of the different baseline models.

Table 12: Descriptions of different baseline models

Model	Description
Fixed number of lags	
Fixed model 1	4 weeks used as a fixed number of lags in estimating baseline sales
Fixed model 2	6 weeks used as a fixed number of lags in estimating baseline sales
Dynamic lag and lead values	
Dynamic model 1	Lags and leads defined based on seasonality and turnover. High-low categorization based on medians, which then drives the lag and lead values for each category
Dynamic model 2	Similar to dynamic model 1 with exception that this model uses fewer weeks
Dynamic model 3	Similar to other dynamic models except for categorization that is extended to contain high-medium-low grouping based on 25th and 75th percentiles

The same analysis steps were followed regarding each alternative as with the initial model. First, the baseline model was run and gross lifts were calculated in relation to each promotion. Then the data was de-seasonalised after which switching and stockpiling models were fitted and results saved. Table 13 puts together an overview of empirical results based on the different baseline models.

Table 13: Comparison of empirical findings based on different baseline models

Model	Total number of categories(1)	Baseline robustness(2)	Robustness p-value(3)	Median switching %	% of significant switching(4)	% of significant stockpiling(5)
Fixed number of lags						
Fixed model 1	150	Pass	0.904	38.5%	92.7%	12.0%
Fixed model 2	146	Pass	0.479	38.3%	93.8%	10.3%
Dynamic lag and lead values						
Dynamic model 1	151	Fail	0.001	35.8%	95.4%	13.2%
Dynamic model 2	151	Fail	0.003	35.8%	96.0%	14.6%
Dynamic model 3	149	Fail	0.000	35.6%	96.6%	13.4%
			Max	***38.5%***	***96.6%***	***14.6%***
			Min	***35.6%***	***92.7%***	***10.3%***
			Average	***36.8%***	***94.9%***	***12.7%***

(1) Categories with less than 30 promotional weeks excluded from the sample because regression results may not be reliable with such a little sample sizes. Moving average used to calculate baselines so number of lags and leads affects how many observations are lost in the beginning and in the end of time series. Hence, total number of categories may vary among the models

(2) Baseline robustness tested by calculating baseline estimates for non-promotional weeks and comparing these to actual unit sales. A paired t-test conducted to test whether the baselines are statistically different to actuals

(3) Null hypothesis: true difference in means equals to 0

(4) % of siginicant coefficient estimates at 1% significance level

(5) % of significant F-values at 1% significance level

First, one should note that the number of categories may change between the models. This stems from the fact that categories with less than 30 promotional weeks were excluded from the sample because regression results may not be reliable with such a little sample sizes. Moving average was used to calculate baselines so the number of lags and leads affects how many observations are lost in the beginning and in the end of time series. Consequently, total number of categories varies slightly among the models.

The results obtained from the validation analysis is promising. The median switching effect does not vary much among the models settling between 35.6% and 38.5%. Also, over 90% of categories have statistically significant switching effects in each scenario. Low level of variation in estimated switching effects combined with high proportion of statistically significant models indicate consistent and reliable results. The stockpiling results do not change much either when altering the baseline model. Similar to the initial analysis the individual coefficient estimates exhibit inconsistent signs in other scenarios as well. Only 10-15% of the overall models measured by the F-test are statistically significant. These findings further support the argument that stockpiling effects may not be present in the sample. Another possible explanation was that the effect sizes can be so small that the applied model relying on POS data cannot capture these effects properly. From the above, one may conclude that the quantitative analysis seems to provide consistent results.

The findings reported in this thesis seem reliable as evidenced by robustness of the models. Moreover, comparability to prior studies further strengthens the credibility of reported findings. So, what can be concluded from this thesis regarding promotions? First, analysing effectiveness of promotions is not easy despite being rather straightforward to go about once the process is known. But applying a systematic approach using quantitative data can provide many benefits to retailers. The kind of holistic and systematic approach introduced in this thesis gives a unified framework to study all promotions. One can evaluate the effects broadly across thousands of products in hundreds of categories following the same method and using the same data. This is often very valuable in practice. Also, the analysis can be easily replicated, revised or even automated as new data is accumulated, thus allowing the retailers to stay up-to-date with relatively small effort.

The overall approach forms the bedrock for evaluating promotion impact, but the parts of the analysis may be interesting and beneficial for retailers on their own. For example, measuring gross lift quantifies the promotion response in a very simple and understandable way. Even if none of the subsequent steps would be involved one can identify the least effective promotions rather quickly. Thus, retailers should always focus on quantifying the gross lift at first to capture the low hanging fruits. That is, cancelling out promotions that are not able to generate sales lift in the first place. Depending on the company objectives, it may vary how much sales lift is required at minimum so that the company is willing to invest in promotions. If promotion does not generate a lift in sales at all or does not meet the criteria set by the company, it does make sense to spend time analysing the effects further. In some cases, switching effects alone give interesting information and helps to understand the dynamics of different product categories. But the main argument here is that the approach becomes the most powerful when individual parts are combined as each step in the analysis builds upon each other. First, one can estimate the baselines and compare these to actual sales to quantify gross lift. Like discussed, using this information one may already cut off least effective promotions. Then switching effects can be added to the analysis which combined with the gross lift calculations allows to quantify short-term incremental lift for a retailer. Further, retailers may include long-term effects such as stockpiling to the analysis or add cost information to calculate promotion margins. The major benefit of this kind of systematic approach is that one can stop at a level that fits the company objectives. For some, measuring short-term incremental lift may suffice whereas others may be more interested in factoring long-term impact of promotions as well.

Another major implication for retailers is that the approach introduced in this paper does not restrict the company stick with single method that cannot be altered later. The individual parts can be modelled differently without affecting the remainder of the analysis giving the modeller freedom to choose the best methods for each part. For example, one could use different approach to model the baselines. Instead of using moving average, which was applied in this thesis, one could use week-of-year averages or even a sophisticated software that considers hundreds of variables to estimate the baseline sales. Similarly, switching or stockpiling models could be altered without affecting the baseline calculations.

5 Conclusions, limitations and future research

5.1 Summary of key findings

The ultimate objective of this thesis was to extract information from a large set of data to improve promotional decision making. The concrete aim of the study was to analyse impact of sales promotions on unit sales and to identify the most and the least effective promotions. Promotion effects were studied from a retailer's perspective that had received attention in prior research to a lesser extent. The research problem was broken down into three specific research questions. A systematic approach was applied to provide answers for each using quantitative data.

The first research question asks whether promotions generate a lift in sales during the promotional period. I studied over 140,000 promotions and found that about 75% of promotions do generate a sales lift. The gross lift i.e. the number of additional units gained due to a promotion was reported as a percentage of the baseline sales. A median gross lift of approximately 130% was found. This means that for every unit of baseline sales 1.3 additional units are gained due to promotion.

However, measuring gross lift alone does not suffice because consumption may switch from other products in the category thus cannibalizing sales of the non-promoted items. This is often an undesirable outcome for retailers that need to account for the whole category sales as opposed to manufacturers that are only interested in their own brand sales. The second research question addressed this problem by asking to what extent is the sales lift switched from other non-promoted products within a category. Switching effects were studied across 150 categories for which there were enough promotional activity to fit the regression model. The findings suggest that approximately 40% of the sales promotion bump is switched from competing products within the category. This means that if a retailer gains 100 units due to promotion, only 60 of them are incremental after accounting for the switching effects in that category.

Apart from modelling the short-term impact of promotions, stockpiling effects were studied which refers to a phenomenon where quantity sold in the weeks after the promotion is decreased. “Do promotions affect future sales via stockpiling?” formed the third research questions to address this problem. Similar regression-based approach to switching was applied and stockpiling effects were studied in the same 150 categories. However, no statistically significant stockpiling effects were found. Prior research suggest that this may happen due to two reasons. First, stockpiling effects are rarely found with aggregate POS data, especially if the effect sizes are believed to be relatively small. This suggests that household data may be needed to capture stockpiling effects. Another reason for the absence of stockpiling effects is that these effects may apply only to certain products or customers implying that stockpiling effects are not present in all categories. Therefore, studies that have included only few categories may have overstated the prevalence of stockpiling.

The results imply that not all promotions generate the desired results. This means that retailers should focus their promotional spending to products and categories that are more effective and generate positive impact. Systematic analysis based on quantitative data helps to understand the true performance and learn from past while guiding effective promotional decisions in the future which oftentimes are a necessity for many companies to keep ahead of competition. Depending on the company objectives and internal analytics capabilities the analysis process for evaluating promotion effectiveness can be adjusted accordingly.

5.2 Limitations and suggestions for future research

Although the dataset in this thesis covered sales data from a two-year period, all categories did not have enough observations (i.e. promotional weeks) to estimate promotion effects. Categories with less than 30 promotional weeks were excluded from the regression analysis due to small sample size. Hence, a longer data set is required if one needs to get estimates for categories with little promotional activity. However, one should keep in mind that at the same time the computational burden increases. The two-year data set used in this thesis consisted of over 4 million observations. This already possessed some limitations to the data processing and required the use of R packages that are optimized for large amounts of data.

In theory, one would like to model a “normal” promotion impact which means that one identifies and deletes known unusual events occurred in the market as well as holiday weeks that may affect the results. Also, one could eliminate store or market areas that are newly entered or do not represent typical store for the company. Considering that this thesis was performed by a single student incorporating large volume of data, the research process is naturally limited to some extent as manual adjustments to the data are limited. If a larger research team were formed and various stakeholders from the retailer company were engaged, one could most likely improve the accuracy and consistency of results by making these kinds of additional adjustments to data.

This study focused on quantifying the promotion effects but did not aim to explain why some promotions are more effective than the others. One interesting avenue for future research thus would be to study key drivers behind the promotion impact. Promotions differ in terms of type and “depth”, brand characteristics, category characteristics and store characteristics. Research on how promotion impact varies with some of these parameters would give additional perspective to the research topic.

One could also experiment with different methods or data for modelling the effects. For example, using different approach than moving average to estimate baselines can provide interesting insights about the magnitude of gross lift. On the other hand, household-level data could be combined with the existing POS data to evaluate more accurately stockpiling effects in consumer purchasing behaviour. However, this would either require a loyalty program to be in place or panel data to be collected by a market research company. It would be interesting to see whether these kinds of alterations to the analysis provided divergent results.

Another research avenue that surely captures the interest of both academics and practitioners’ is the profit impact of promotion. In theory, the analysis concerning promotion impact on unit sales followed in this thesis could be extended to profit issues. However, measuring profit impact most likely requires significant resources from the company and from the research team including rigorous allocation of various trade allowances from manufactures and combining data from several sources, among other issues. This calls for high ambition levels from the retailer as well as willingness to invest substantial time and resources to the process while engaging stakeholders from different parts of the organization. Furthermore, a major barrier in this analysis is the complexity in calculating the promotional margins accurately because all costs need to be allocated for each individual promotion. Nevertheless, research on profit impact serves as another attractive direction for future studies.

References

Abraham, M. M., & Lodish, L. M. (1993). An implemented system for improving promotion productivity using store scanner data. *Marketing Science, 12*(3), 248-269.

Ailawadi, K., & Gupta, S. (2014). Sales promotions. *History of Marketing Science, Singapore,* , 463-497.

Ailawadi, K., Harlam, B. A., Cesar, J., & Trounce, D. (2006). Promotion profitability for a retailer: The role of promotion, brand, category, and store characteristics. *Journal of Marketing Research, 43*(4), 518-535.

Ailawadi, K., Harlam, B. A., César, J., & Trounce, D. (2007). Practice prize Report—Quantifying and improving promotion effectiveness at CVS. *Marketing Science, 26*(4), 566-575.

Barton, D., & Court, D. (2012). Making advanced analytics work for you. *Harvard Business Review, 90*(10), 78-83.

Bell, D. R., Chiang, J., & Padmanabhan, V. (1999). The decomposition of promotional response: An empirical generalization. *Marketing Science, 18*(4), 504-526.

Blattberg, R. C., & Briesch, R. (2012). Sales promotions. In Ö. Özalp & R. Phillips (Eds.), *The oxford handbook of pricing management* (First ed., pp. 585-619) Oxford University Press.

Blattberg, R. C., Briesch, R., & Fox, E. J. (1995). How promotions work. *Marketing Science, 14*(3_supplement), G122-G132.

Blattberg, R. C., & Neslin, S. A. (1990). *Sales promotion: Concepts, methods, and strategies* Prentice Hall.

Boston Consulting Group. (2014). Making big data work: Retailing. Retrieved August 8, 2017 from https://www.bcgperspectives.com/content/articles/retail_pricing_making_big_data_work_retailing/

Boston Consulting Group. (2015). How retailers can improve promotion effectiveness. Retrieved August 14, 2017 from https://www.bcgperspectives.com/content/articles/retail-pricing-how-retailers-can-improve-promotion-effectiveness/

Boston Consulting Group. (2017). Improving promotional effectiveness. Retrieved August 18, 2017 from https://www.bcg.com/expertise/industries/retail/improving-promotional-effectiveness.aspx

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: How does data-driven decisionmaking affect firm performance?

Bucklin, R. E., Gupta, S., & Siddarth, S. (1998). Determining segmentation in sales response across consumer purchase behaviors. *Journal of Marketing Research,* , 189-197.

Bucklin, R. E., & Lattin, J. M. (1992). A model of product category competition among grocery retailers. *Journal of Retailing, 68*(3), 271.

Chiang, J. (1991). A simultaneous approach to the whether, what and how much to buy questions. *Marketing Science, 10*(4), 297-315.

Chintagunta, P. K. (1993). Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science, 12*(2), 184-208.

CrossCap. (2017). 3 strategies for measuring sales lift in a retail promotion. Retrieved August 28, 2017 from https://www.crosscap.com/retail-marketing-blog/3-strategies-for-measuring-sales-lift-in-a-retail-promotion

Davenport, T. H. (2017). How analytics has changed in the last 10 years (and how it's stayed the same). *Harvard Business Review,* , 28/08/2017.

Davenport, T. H., Harris, J. G., & Morison, R. (2010). *Analytics at work: Smarter decisions, better results* Harvard Business Press.

Dipl-Kfm, J. A. K. (2015). Sales promotion management in retailing: Tasks, benchmarks, and future trends. *Marketing Review St.Gallen, 32*(3), 67.

Fader, P. S., & Hardie, B. G. (1996). Modeling consumer choice among SKUs. *Journal of Marketing Research,* , 442-452.

Gaudiano, P. (2017). The best approach to decision making combines data and managers expertise. *Harvard Business Review,* , 28.08.2017.

Gedenk, K., Neslin, S. A., & Ailawadi, K. L. (2010). Sales promotion. *Retailing in the 21st century* (pp. 393-407) Springer.

Guadagni, P. M., & Little, J. D. (1983). A logit model of brand choice calibrated on scanner data. *Marketing Science, 2*(3), 203-238.

Gupta, S. (1988). Impact of sales promotions on when, what, and how much to buy. *Journal of Marketing Research,* , 342-355.

Hendel, I., & Nevo, A. (2003). The post-promotion dip puzzle: What do the data have to say? *Quantitative Marketing and Economics, 1*(4), 409-424.

Kotler, P., & Armstrong, G. (2014). *Principles of marketing* Pearson education.

Macé, S., & Neslin, S. A. (2004). The determinants of pre-and postpromotion dips in sales of frequently purchased goods. *Journal of Marketing Research, 41*(3), 339-350.

Manchanda, P., Ansari, A., & Gupta, S. (1999). The “shopping basket”: A model for multicategory purchase incidence decisions. *Marketing Science, 18*(2), 95-114.

McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review, 90*(10), 60-68.

McKinsey&Company. (2013a). Applying advanced analytics in consumer companies. Retrieved August 25, 2017 from http://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/applying-advanced-analytics-in-consumer-companies

McKinsey&Company. (2013b). The power of advanced analytics in revenue management. Retrieved August 18, 2017 from http://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/retail/articles/perspectives%20-%20winter%202013/power_of_advanced_analytics_in_revenue_management_periscope%20vf.ashx

Nair, H., Dubé, J., & Chintagunta, P. (2005). Accounting for primary and secondary demand effects with aggregate data. *Marketing Science, 24*(3), 444-460.

Narasimhan, C., Neslin, S. A., & Sen, S. K. (1996). Promotional elasticities and category characteristics. *The Journal of Marketing,* , 17-30.

Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking* " O'Reilly Media, Inc.".

Shearer, C. (2000). The CRISP-DM model: The new blueprint for data mining. *Journal of Data Warehousing, 5*(4), 13-22.

Sun, B. (2005). Promotion effect on endogenous consumption. *Marketing Science, 24*(3), 430-443.

Van Heerde, H. J. (2005). The proper interpretation of sales promotion effects: Supplement elasticities with absolute sales effects. *Applied Stochastic Models in Business and Industry, 21*(4-5), 397-402.

Van Heerde, H. J., Gupta, S., & Wittink, D. R. (2003). Is 75% of the sales promotion bump due to brand switching? no, only 33% is. *Journal of Marketing Research, 40*(4), 481-491.

Van Heerde, H. J., Leeflang, P. S., & Wittink, D. R. (2000). The estimation of pre-and postpromotion dips with store-level scanner data. *Journal of Marketing Research, 37*(3), 383-395.

Van Heerde, H. J., Leeflang, P. S., & Wittink, D. R. (2004). Decomposing the sales promotion bump with store data. *Marketing Science, 23*(3), 317-334.

Van Heerde, H. J., & Neslin, S. A. (2017). Sales promotion models. *Handbook of marketing decision models* (pp. 13-77) Springer.